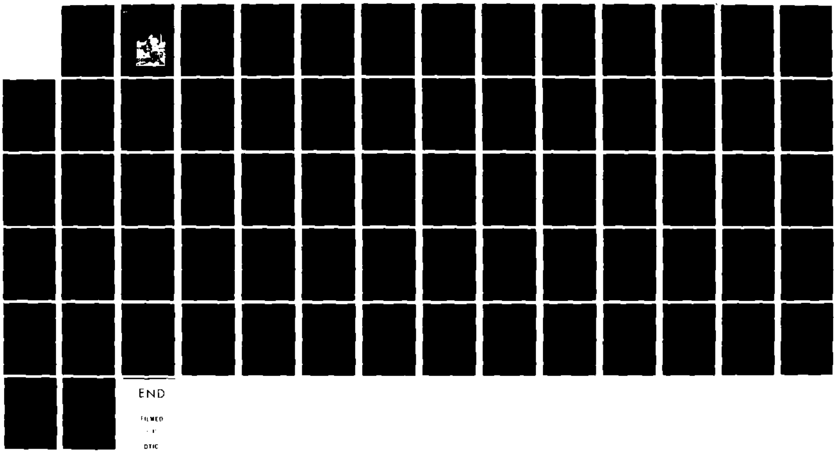
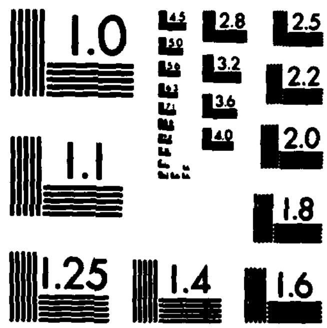
AD-A120 972 SIGNAL DETECTION FOR PARETO RENEWAL PROCESSES(U) SAN DIEGO STATE UNIV CA DEPT OF MATHEMATICAL SCIENCES C B BELL ET AL. OCT 82 8-82 N00014-80-C-0208

UNCLASSIFIED F/G 12/1 NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

(2)

SDSU

8-82

SERIES IN STATISTICS AND BIOSTATISTICS

SIGNAL DETECTION FOR PARETO RENEWAL PROCESSES

C.B. BELL, R. AHMAD, C.J. PARK, R. LUI

San Diego State University
Department of Mathematical Sciences

SIGNAL DETECTION FOR PARETO RENEWAL PROCESSES

C. B. Bell*, R. Ahmad**, C. J. Park*, R. Lui*

*San Diego State University

**University of Strathclyde, Scotland

1. Introduction and Summary

Several types of applications entail point processes with "dead" times after each event. One such family of stochastic processes is Ω(PRP), the family of Pareto Renewal Processes. The i.i.d. interarrival times $\{X_j\}$ satisfy, $F(x) = P\{X \leq x\} = 1 - (\frac{A}{x})^s$, $x > A > 0$ and $s > 0$.

An additional interesting property of the interarrival-time distributions is that they are all "thick-tailed" relative to the corresponding distributions for Gaussian processes and Poisson processes. Further, a variety of tail thicknesses, one for each s-value, is possible. These two properties lead to some interesting inference problems, of which one is here concerned only with signal detection.

[The Pareto distribution itself was, of course, introduced by Vilfredo Pareto (1848 - 1923). (See Reference [22]). This distribution has been used and studied by numerous other authors including Pigon (1932)

This work was principally supported by the Office of Naval Research through Grant No. N00014-80C-0208.

Mandelbrot (1960, 1963), Fisk (1961) and Johnson and Kotz (1970).]

Both one-sample and two-sample signal detection problems with historical data will be considered here. The organization of the paper is as follows. Section 2 contains the basic properties of the inter-arrival-time distributions and the MLE's (maximum likelihood estimates) under various circumstances. Section 3 contains the distribution theory necessary for insight and inference. Most of the proofs for the results of this section are straightforward. In Section 4, one introduces the fundamental statistical concepts to be used, namely, (i) the BDT (Basic Data Transformation); (ii) the M-S-S (minimal sufficient statistic); (iii) the M-S-N (maximal statistical noise); (iv) PDF (parametric distribution-free) statistics; and (v) NPDF (nonparametric distribution-free) statistics.

Section 5 discusses the uses of the Kolmogorov-Smirnov (1933) statistic; and its extensions by Lilliefors (1967, 1969), Srinivasan (1970) and Y. Choi (1980) in signal detection. Sections 6 and 7 treat the one-sample and two-sample detection problems, respectively. Tables summarizing the results, are presented at the end of appropriate sections.

Appendix A contains proofs of some of the assertions made in the paper. Appendix B contains numerical examples illustrating each of the detection procedures derived in Sections 6 and 7.

2. Elementary Properties and Estimators of Pareto Renewal Process

Most of the results of this section are adaptations of results in Johnson and Kotz (1970). The Pareto distribution $Pa(A,s)$ is defined for each $A > 0$, $s > 0$ by

$$F(x) = \begin{cases} 1 - (\frac{A}{x})^s & x > A \\ 0 & x \leq A \end{cases}$$

This is a special form of the Pearson type VI distribution. The Pareto density function is

$$f(x) = sA^s \, X^{-s-1} \quad \text{for} \quad x > A > 0.$$

If $X \sim Pa(A,s)$, then the rth moment of X exists if and only if $r < s$ and is given by

$$E(X^r) = \frac{sA^r}{s - r}.$$

The variance of X is $sA^2(s-1)^{-2}(s-2)^{-1}$ for $s > 2$. For further details on moments, see Malik (1966).

Malik (1966, 1967) has also obtained results on the characteristic functions of the order statistics from a Pareto distribution as well as recurrence relations between the moments and covariance of the order statistics. Levy (1925) discovered a class of stable laws (Stable Paresian) which follow the asymptotic form of the Pareto law.

If $X \sim Pa(A,s)$, $Y = \ln X$, then $Y \sim T\text{-}\exp(\ln A,s)$ where

T-exp(B,s) is truncated-exponential with distribution function $H(x) = 1 - e^{-s(x - B)}$ for $x \geq B$. Also if $Z = X^{-s}$, then $Z \sim U(0, A^{-s})$. Inference for the truncated-exponential has been studied by Park and U. Choi (1978), Beg (1981) and signal detection problem for the uniform process has been investigated by Y. Choi-Bell-Ahmad-Park (1982). Some of those results will be used and compared to the ones herein.

Here one is primarily interested in MLE's of the parameters A and s. Other types of estimates are given in Johnson and Kotz (1970). The likelihood function for the interarrival times $(X_1, X_2, \ldots, X_n)$ from a PRP is $L = \prod_{j=1}^{n} \frac{sA^s}{X_j^{s+1}}$. The proofs of the following theorems are straightforward.

<u>Theorem 2.1</u>. (One-parameter, $s = s_0$ known)

(i) The MLE of A is $\hat{A} = X(1)$, which is distributed $Pa(A, ns_0)$;

(ii) $E(\hat{A}) = ns_0A(ns_0 - 1)^{-1}$ for $ns_0 > 1$ and $V(\hat{A}) = ns_0A^2[(ns_0 - 2)(ns_0 - 1)^2]^{-1}$ for $ns_0 > 2$;

(iii) The MVUE (minimum variance unbiased estimator) of A is $A^* = (ns_0 - 1)(ns_0)^{-1}X(1)$ for $ns_0 > 1$, and $V(A^*) = [ns_0(ns_0 - 2)]^{-1}A^2$ for $ns_0 > 2$,

(iv) Both $\hat{A}$ and A^* are consistent estimators of A, i.e., $\hat{A}/A \to 1$, $A^*/A \to 1$ as $n \to \infty$ but only A^* is unbiased.

(v) the M-S-S (minimal sufficient statistic) for A is X(1) and the family $\{Pa(A, ns_0)\}$ indexed by $A > 0$ is complete.

<u>Theorem 2.2</u>. (One-parameter, $A = A_0$ known)

(i) The MLE of s is $\hat{s} = \dfrac{n}{\sum_{j=1}^{n} \ln X_j - n \ln A_0}$;

(ii) $\dfrac{2ns}{\hat{s}} \sim \chi^2_{2n}$, $E(\frac{1}{\hat{s}}) = \frac{1}{s}$ and $V(\frac{1}{\hat{s}}) = \frac{1}{ns^2}$;

(iii) $E(\hat{s}) = (\frac{n}{n-1})s$ for $n > 1$ and $V(\hat{s}) = \dfrac{n^2 s^2}{(n-1)^2(n-2)}$ for $n > 2$;

(iv) The MVUE of s is given by $s^* = (\frac{n-1}{n})\hat{s}$ and $V(s^*) = \dfrac{s^2}{n-2}$ for $n > 2$;

(v) Both $\hat{s}$ and s^* are consistent estimators of s but only s^* is unbiased; and

(vi) the M-S-S for s is $\hat{s}$ (or s^*).

<u>Theorem 2.3</u>. (Two parameters, A, s both unknown)

(i) The MLE of (A,s) is $(\hat{A},\hat{s})$ where

$$\hat{A} = X(1) \quad \text{and} \quad \hat{s} = \frac{n}{\sum_{j=1}^{n} \ln X_j - n\ln X(1)};$$

(ii) $\hat{A}$ and $\hat{s}$ are independent, $\hat{A} \sim Pa(A,ns)$, $\dfrac{2ns}{\hat{s}} \sim \chi^2_{2(n-1)}$;

(iii) $E(\hat{A}) = \dfrac{nsA}{ns-1}$ for $ns > 1$, $V(\hat{A}) = \dfrac{nsA^2}{(ns-2)(ns-1)^2}$ for $ns > 2$;

(iv) $E(\hat{s}) = \dfrac{ns}{n-2}$ for $n > 2$, $V(\hat{s}) = \dfrac{n^2 s^2}{(n-2)^2(n-3)}$ for $n > 3$;

(v) The MVUE of A is $A^* = \dfrac{(ns-1)}{ns} X(1)$ for $ns > 1$ and $V(A^*) = \dfrac{A^2}{ns(ns-2)}$ for $ns > 2$;

(vi) the MVUE of s is $s^* = \frac{(n-2)\hat{s}}{n}$ and $V(s^*) = \frac{s^2}{n-3}$;

(vii) the M-S-S for (A,s) is $(\hat{A},\hat{s})$.

The results of this section are summarized in Table 2.1 below. The last column of this table yields the M-S-N (maximal statistical noise) for various detection problems. A formal definition of the M-S-N is given in Section 4. As far as Table 2.1 is concerned, one should view the M-S-N, $N(\underset{\sim}{x})$, as complementary to and statistically independent of the M-S-S, $S(\underset{\sim}{x})$. Several versions of the M-S-N are given in Section 3.

3. Distribution Theory for Pareto Renewal Processes

In this section one develops some results which yield the M-S-S (to be defined in Section 4) for various one-sample and two-sample problems.

Let $X_1, X_2, \ldots, X_n$ be i.i.d. $Pa(A,s)$ and let $X(1) \leq X(2) \leq \ldots \leq X(n)$ denote the order statistics of the X's.

Definition 3.1. One denotes by G-O-S(k) the distribution of the order statistics induced by a random sample of size k drawn from a distribution $G(\cdot)$. The following lemma is fundamental.

Lemma 3.1. (Renyi (1953), Pyke (1965)). Let $\xi_1, \xi_2, \ldots, \xi_k$ be

TABLE 2.1.

	MLE	Dist of MLE	E(MLE)	Var(MLE)	MVUE(·)	Var(MVUE)	MSS(·)	**(MSN(·))
1-parameter A unknown $s = s_0$ known Theorem 2.1.	of A is $\hat{A} = X(1)$	$Pa(A, ms_0)$	$\frac{ms_0A}{ms_0-1}$ for $ms_0 > 1$	$\frac{ms_0A^2}{(ms_0-2)(ms_0-1)^2}$ for $ms_0 > 2$	of A is $A^* = (\frac{ms_0-1}{ms_0})X(1)$ for $ms_0 > 1$	$\frac{A^2}{ms_0(ms_0-2)}$ for $ms_0 > 2$	of A is $X(1)$	$(Q_1, \ldots, Q_{n-1}, \underset{\sim}{R})$ $\underset{\sim}{R}$ = rank vector $\{Q_j\}$ may be anyone of the following: $\{U_j\}$, $\{V_j\}$, $\{W_j\}$, $\{Y_j\}$, $\{Z_j\}$
1-parameter s unknown $A = A_0$ known Theorem 2.2.	of s is $\hat{s} = \frac{n}{\sum_{j=1}^{n} \ln X_j - n\ln A_0}$	$\frac{2ns}{\hat{s}} \sim \chi^2_{2n}$	$(\frac{n}{n-1})s$ for $n > 1$	$\frac{n^2s^2}{(n-1)^2(n-2)}$ for $n > 2$	of s is $s^* = (\frac{n-1}{n})\hat{s}$ for $n > 1$	$\frac{s^2}{n-2}$ for $n > 2$	of s is $\hat{s}$	$(\frac{Y_1}{Y_n}, \frac{Y_2}{Y_n}, \ldots, \frac{Y_{n-1}}{Y_n}) \sim$ U-O-S(n-1)
2-parameters (A,s) unknown Theorem 2.3.	of (A,s) is $\hat{A} = X(1)$ $\hat{s} = \frac{n}{\sum_{j=1}^{n} \ln X_j - n\ln X(1)}$	$\hat{A} \sim Pa(A, ms)$ $\frac{2ns}{\hat{s}} \sim \chi^2_{2n-2}$	$E(\hat{A}) = \frac{msA}{ms-1}$ $E(\hat{s}) = \frac{ns}{n-2}$	$Var(\hat{A}) = \frac{msA^2}{(ms-2)(ms-1)^2}$ $Var(\hat{s}) = \frac{n^2s^2}{(n-2)^2(n-3)}$	of A is $A^* = \frac{(ms-1)X(1)}{ms}$ of s is $s^* = \frac{(n-2)\hat{s}}{n}$	$Var(A^*) = \frac{A^2}{ms(ms-2)}$ $Var(s^*) = \frac{s^2}{n-3}$	of (A,s) is $(\hat{A}, \hat{s})$	$(\frac{\beta_1}{\beta_{n-1}}, \ldots, \frac{\beta_{n-2}}{\beta_{n-1}}, R(X_1), \ldots, R(X_n))$, where $R(X_i)$ = rank of X_i

* $\{U_j\}$, $\{V_j\}$, $\{W_j\}$, $\{Y_j\}$, $\{Z_j\}$, $\{\beta_j\}$, $\underset{\sim}{R}$, $\{Q_j\}$, $\{R_j\}$ are defined in Section 3, Definition 3.2.

** M-S-N is defined in Section 4.

i.i.d. Exp(λ). Then (i) $k\bar{\xi} = \sum_{j=1}^{k} \xi_j \sim \Gamma(k,\lambda)$, (ii) $\underset{\sim}{\xi}^* =$

$$\left\{ \frac{\xi_1}{k\bar{\xi}}, \frac{\xi_2}{k\bar{\xi}}, \ldots, \frac{\sum_{j=1}^{k-1} \xi_j}{k\bar{\xi}} \right\} \sim$$ U-O-S$(k - 1)$ where $U(\cdot)$ is the $U(0,1)$ distribution, (iii) $k\bar{\xi}$ and $\underset{\sim}{\xi}^*$ are independent.

<u>Remark</u>. Note that λ has been eliminated in $\underset{\sim}{\xi}^*$.

<u>Lemma 3.2</u>. Let $\eta_1, \eta_2, \ldots, \eta_k$ be Exp(λ) - O - S(k). Define $\eta_1^*, \ldots, \eta_k^*$ by $\eta_1^* = k\eta_1$, $\eta_2^* = \eta_1 + (k - 1)\eta_2, \ldots, \eta_j^* = \eta_1 + \ldots + \eta_{j-1} + (k + 1 - j)\eta_j, \ldots, \eta_k^* = k\bar{\eta}$ and $\underset{\sim}{\eta}^{**} =$

$$\left\{ \frac{\eta_1^*}{\eta_k^*}, \frac{\eta_2^*}{\eta_k^*}, \ldots, \frac{\eta_{k-1}^*}{\eta_k^*} \right\}.$$ Then

(i) $\eta_k^* \sim \Gamma(k,\lambda)$ is the M-S-S for λ;

(ii) $\underset{\sim}{\eta}^{**} \sim$ U-O-S$(k - 1)$;

(iii) η_k^* and $\underset{\sim}{\eta}^{**}$ are independent.

<u>Theorem 3.3</u>. Let $X_1, X_2, \ldots, X_n$ be i.i.d. Pa(A,s), then

(i) $\{\ln \frac{X_1}{A}, \ln \frac{X_2}{A}, \ldots, \ln \frac{X_n}{A}\}$ i.i.d. Exp(s);

(ii) $\{\ln X(1), \ln X(2), \ldots, \ln X(n)\} \sim$ T - Exp($\ln A,s$) - O - S(n);

(iii) $\{\ln \frac{X(1)}{A}, \ln \frac{X(2)}{A}, \ldots, \ln \frac{X(n)}{A}\} \sim$ Exp(s) - O - S(n);

(iv) $\{\frac{A}{X(n)}, \frac{A}{X(n-1)}, \ldots, \frac{A}{X(1)}\} \sim P_w(s)$-O-S(n) where $P_w(s)$ is the power distribution defined by $H(x) = \begin{cases} 0 & x \le 0 \\ x^s & 0 < x < 1, \\ 1 & x \ge 1 \end{cases}$

(v) $\{[\frac{1}{X(n)}]^s, [\frac{1}{X(n-1)}]^s, \ldots, [\frac{1}{X(1)}]^s\} \sim U(0,A^{-s})$-O-S(n),

(vi) $\{[\frac{A}{X(n)}]^s, [\frac{A}{X(n-1)}]^s, \ldots, [\frac{A}{X(1)}]^s\} \sim$ U-O-S(n).

<u>Definition 3.2</u>. Let $X_1, X_2, \ldots, X_n$ be i.i.d. Pa(A,s). [The following random variables T_r, D_r, E_r, $1 \le r \le n-1$ will be used frequently throughout the rest of the paper and are mentioned in Table 2.1.]

(i) $T_r = \sum_{j=1}^{r} \ln(\frac{X_j}{A})$, $1 \le r \le n$;

(ii) $D_r = \sum_{j=1}^{r-1} \ln(\frac{X(j+1)}{X(1)}) + (n-r) \ln \frac{X(r+1)}{X(1)}$, $1 \le r \le n-1$;

(iii) $E_r = \sum_{j=1}^{r} (n+1-j) \ln(\frac{X(j+1)}{X(j)})$, $1 \le r \le n-1$.

The following theorem is an easy consequence of Lemmas 3.1 and 3.2.

<u>Theorem 3.4</u>. (i) $\underset{\sim}{T} = \{\frac{T_1}{T_n}, \frac{T_2}{T_n}, \ldots, \frac{T_{n-1}}{T_n}\} \sim U$-O-S(n - 1),

and is independent of $T_n \sim \Gamma(n,s)$;

(ii) $\underset{\sim}{D} = \left(\frac{D_1}{D_{n-1}}, \frac{D_2}{D_{n-1}}, \ldots, \frac{D_{n-2}}{D_{n-1}} \right) \sim$ U-O-S(n - 2) and is independent of $D_{n-1} \sim \Gamma(n-1,s)$;

(iii) $\underset{\sim}{E} = \left(\frac{E_1}{E_{n-1}}, \frac{E_2}{E_{n-2}}, \ldots, \frac{E_{n-2}}{E_{n-1}} \right) \sim$ U-O-S(n - 2) and is independent of $E_{n-1} \sim \Gamma(n-1,s)$.

The following theorem contains various versions of the M-S-N (see Definition 4.1) when A is unknown. They will be used in the Kolmogorov-Smirnov statistics for the one-sample case in Section 6.

<u>Theorem 3.5</u>. (i) $\left(\ln \left[\frac{X(2)}{X(1)}\right], \ln\left[\frac{X(3)}{X(1)}\right], \ldots, \ln\left[\frac{X(n)}{X(1)}\right] \right) \sim$ Exp(s)-O-S(n - 1), (ii) $\left(\frac{X(2)}{X(1)}, \frac{X(3)}{X(1)}, \ldots, \frac{X(n)}{X(1)} \right) \sim$ Pa(1,s)-O-S(n - 1);

(iii) $\left(\left[\frac{X(1)}{X(n)}\right]^s, \left[\frac{X(1)}{X(n-1)}\right]^s, \ldots, \left[\frac{X(1)}{X(2)}\right]^s \right) \sim$ U-O-S(n - 1);

(iv) $\left(\left[\frac{X(1)}{X(2)}\right]^{(n-1)s}, \left[\frac{X(2)}{X(3)}\right]^{(n-2)s}, \ldots, \left[\frac{X(n-1)}{X(n)}\right]^s \right)$ i.i.d. U(0,1);

(v) $\left(n \ln \left[\frac{X(2)}{X(1)}\right], (n-1)\ln\left[\frac{X(3)}{X(2)}\right], \ldots, \ln\left[\frac{X(n)}{X(n-1)}\right] \right)$ i.i.d. Exp(s).

<u>Definition 3.3</u>. For $1 \le j \le n-1$, one defines

(i) $V_j = \ln\left[\frac{X(j+1)}{X(1)}\right]$, (ii) $U_j = \left[\frac{X(j+1)}{X(1)}\right]$, (iii) $W_j = \left[\frac{X(1)}{X(n-j+1)}\right]^s$,

(i $Y_j = \left[\frac{X(j)}{X(j+1)}\right]^{(n-j)s}$, (v) $Z_j = (n+1-j)\ln\left[\frac{X(j+1)}{X(j)}\right]$.

They will be used throughout the rest of the paper.

The final lemma for this section deals with a two-sample situation.

Lemma 3.6. Let $X_1, X_2, \ldots, X_m$, $Y_1, Y_2, \ldots, Y_n$ be independent with $X_j \sim Pa(A_1,s_1)$ and $Y_j \sim Pa(A_2,s_2)$. Let $\eta_1 = \sum_{j=1}^{m-1} \ln[\frac{X(j+1)}{X(1)}]$,

$\eta_2 = \sum_{j=1}^{n-1} \ln[\frac{Y(j+1)}{Y(1)}]$, then (i) $X(1), Y(1), \eta_1, \eta_2$ are independent; (ii) $X(1) \sim Pa(A_1, ms_1)$, $Y(1) \sim Pa(A_2,ns_2)$; (iii) $\eta_1 \sim \Gamma(m-1,s_1)$, $\eta_2 \sim \Gamma(n-1,s_2)$; (iv) $2ms_1 \ln \frac{X(1)}{A_1}$ and $2ns_2 \ln \frac{Y(1)}{A_2}$ are i.i.d. $\chi_2^2 = Exp(1/2)$; (v) $2s_1\eta_1 \sim \chi^2_{2m-2}$, $2s_2\eta_2 \sim \chi^2_{2n-2}$ and $2s_1\eta_1 + 2s_2\eta_2 \sim \chi^2_{2N-4}$ where $N = m + n$; and

$$\text{(vi)} \quad \frac{(N-2)\left|ms_1 \ln \frac{X(1)}{A_1} - ns_2 \ln \frac{Y(1)}{A_2}\right|}{2s_1\eta_1 + 2s_2\eta_2} \sim F(2, 2N-4).$$

Remark. Lemma 3.6 (vi) is useful when $m = n$ because if $A_1 = A_2$ and $s_1 = s_2$ then the expression on the left does not involve A nor s. [See Section 7.]

The results of this section are summarized in Table 3.1 below.

4. Basic Statistical Concepts

(A) The BDT (Basic Data Transformation) and MSN (Maximal Statistical Noise).

TABLE 3.1. Distribution Theory for Pareto Renewal Processes

Theorem or Lemma	Assumptions	Conclusions
Lemma 3.1.	$\xi_1, \xi_2, \ldots, \xi_k$ i.i.d. $\text{Exp}(\lambda)$	(i) $k\bar{\xi} \sim \Gamma(k,\lambda)$ (ii) $\underset{\sim}{\xi}^* = \{\frac{\xi_1}{k\bar{\xi}}, \frac{\xi_1+\xi_2}{k\bar{\xi}}, \ldots, \frac{\sum_{j=1}^{k-1}\xi_j}{k\bar{\xi}}\}$ $U(0,1)$-O-S$(k-1)$ (iii) $k\bar{\xi}$, $\underset{\sim}{\xi}^*$ are independent (iv) $\underset{\sim}{\xi}^*$ is the M-S-N for λ.
Lemma 3.2.	$\eta_1, \eta_2, \ldots, \eta_k \sim \text{Exp}(\lambda)$-O-S$(k)$, and let $\eta_1^* = k\eta_1^*$ $\eta_1^* = \eta_1+(k-1)\eta_2, \ldots,$ $\eta_j^* = \eta_1+\eta_2+\ldots+\eta_{j-1}+(k-j+1)\eta_j,$ $\eta_k^* = k\bar{\eta}.$	(i) η_k^* is M-S-S for λ, $\eta_k^* \sim \Gamma(k,\lambda)$ (ii) $\underset{\sim}{\eta}^{**} = \{\frac{\eta_1^*}{\eta_k^*}, \ldots, \frac{\eta_{k-1}^*}{\eta_k^*}\}$ and η_k^* are independent (iii) $\underset{\sim}{\eta}^{**} \sim U(0,1)$-O-S$(k-1)$.
Theorem 3.3.	$X_1, X_2, \ldots, X_n$ be i.i.d. $\text{Pa}(A,s)$	(i) $\{\ln\frac{X_1}{A}, \ln\frac{X_2}{A}, \ldots, \ln\frac{X_n}{A}\}$ i.i.d. $\text{Exp}(s)$ (ii) $\{\ln X(1), \ln X(2), \ldots, \ln X(n)\} \sim$ T-Exp$(\ln A, s)$-O-S(n) (iii) $\{\ln\frac{X(1)}{A}, \ln\frac{X(2)}{A}, \ldots, \ln\frac{X(n)}{A}\} \sim \text{Exp}(s)$-O-S$(n)$ (iv) $\{\frac{A}{X(n)}, \frac{A}{X(n-1)}, \ldots, \frac{A}{X(1)}\} \sim P_w(s)$-O-S$(n)$ (v) $\{[\frac{1}{X(n)}]^s, [\frac{1}{X(n-1)}]^s, \ldots, [\frac{1}{X(1)}]^s\} \sim U(0, A^{-s})$-O-S$(n)$ (vi) $\{[\frac{A}{X(n)}]^s, [\frac{A}{X(n-1)}]^s, \ldots, [\frac{A}{X(1)}]^s\} \sim$ U-O-S(n).

Table 3.1. (Continued)

Theorem 3.4.	$X_1, X_2, \ldots, X_n$ i.i.d. Pa(A,s), define (i) $T_r = \sum_{j=1}^{r} \ln \frac{X_j}{A}$ for $1 \le r \le n$ (ii) $D_r = \sum_{j=1}^{r-1} \ln[\frac{X(j+1)}{X(1)}] + (n-r)\ln \frac{X(r+1)}{X(1)}$ for $1 \le r \le n-1$ (iii) $E_r = \sum_{j=1}^{r} (n+1-j)\ln[\frac{X(j+1)}{X(j)}]$ for $1 \le r \le n-1$.	(i) $\underset{\sim}{D} = \{\frac{D_1}{D_{n-1}}, \frac{D_2}{D_{n-1}}, \ldots, \frac{D_{n-2}}{D_{n-1}}\} \sim$ U-O-S(n-2) $\underset{\sim}{D}$ and D_{n-1} are independent and $D_{n-1} \sim \Gamma(n-1,s)$. (ii) $\underset{\sim}{T} = \{\frac{T_1}{T_n}, \frac{T_2}{T_n}, \ldots, \frac{T_{n-1}}{T_n}\} \sim$ U-O-S(n-1) $\underset{\sim}{T}$ and T_n are independent and $T_r \sim \Gamma(r,s)$ for $1 \le r \le n$. (iii) $\underset{\sim}{E} = \{\frac{E_1}{E_{n-1}}, \frac{E_2}{E_{n-1}}, \ldots, \frac{E_{n-2}}{E_{n-1}}\} \sim$ U-O-S(n-2) $\underset{\sim}{E}$ and E_{n-1} are independent and $E_j \sim \Gamma(j,s)$ for $1 \le j \le n-1$.
Theorem 3.5.	$X_1, X_2, \ldots, X_n$ i.i.d. Pa(A,s), for $1 \le j \le n-1$. (i) $V_j = \ln[\frac{X(j+1)}{X(1)}]$, (ii) $U_j = [\frac{X(j+1)}{X(1)}]$ (iii) $W_j = [\frac{X(1)}{X(n-j+1)}]^s$ (iv) $Y_j = [\frac{X(j)}{X(j+1)}]^{(n-j)s}$ (v) $Z_j = (n+1-j)\ln[\frac{X(j+1)}{X(j)}]$	(i) $V_1, V_2, \ldots, V_{n-1} \sim$ Exp(s)-O-S(n-1); (ii) $U_1, U_2, \ldots, U_{n-1} \sim$ Pa(1,s)-O-S(n-1); (iii) $W_1, W_2, \ldots, W_{n-1} \sim$ U-O-S(n-1); (iv) $Y_1, Y_2, \ldots, Y_{n-1} \sim$ i.i.d. U(0,1); (v) $Z_1, Z_2, \ldots, Z_{n-1} \sim$ i.i.d. Exp(s).

Table 3.1. (Continued)

Lemma 3.6.	$X_1, X_2, \ldots, X_m$ i.i.d. $Pa(A_1, s_1)$ independent of $Y_1, Y_2, \ldots, Y_n$ i.i.d. $Pa(A_2, s_2)$ $\eta_1 = \sum_{j=1}^{m-1} \ln[\frac{X(j+1)}{X(1)}]$, $\eta_2 = \sum_{j=1}^{n-1} \ln[\frac{Y(j+1)}{Y(1)}]$	(i) $X(1)$, $Y(1)$, η_1, η_2 are independent; (ii) $X(1) \sim Pa(A_1, ms_1)$, $Y(1) \sim Pa(A_2, ns_2)$; (iii) $\eta_1 \sim \Gamma(m-1, s_1)$, $\eta_2 \sim \Gamma(n-1, s_2)$ (iv) $2ms_1 \ln \frac{X(1)}{A_1}$, $2ns_2 \ln \frac{Y(1)}{A_2}$ are i.i.d. X_2^2; (v) $2s_1\eta_1 \sim X^2_{2m-2}$, $2s_2\eta_2 \sim X^2_{2n-2}$, $2s_1\eta_1 + 2s_2\eta_2 \sim X^2_{2N-4}$; (vi) $(N-2)\left\|ms_1 \ln \frac{X(1)}{A_1} - ns_2 \ln \frac{Y(1)}{A_2}\right\|[2s_1\eta_1 + 2s_2\eta_2]^{-1} \sim F(2, 2N-4)$

Let the generic data point be denoted by $\underset{\sim}{Z} = (X_1, \ldots, X_n)$ in the one-sample cases and by $\underset{\sim}{Z} = (X_1, \ldots, X_m, Y_1, \ldots, Y_n) = (Z_1, \ldots, Z_N)$ in the two-sample cases; and let $S(\underset{\sim}{Z})$ denote the M-S-S.

<u>Definition 4.1</u>. Let $N(\underset{\sim}{Z})$ be a (vector-valued) statistic independent of $S(\underset{\sim}{Z})$ and such that $\delta(\underset{\sim}{Z}) = [S(\underset{\sim}{Z}), N(\underset{\sim}{Z})]$ is 1-1 a.e. Then, (i) $\delta(\cdot)$ is called the BDT; and (ii) $N(\underset{\sim}{Z})$ is called the M-S-N.

It is known that $S(\underset{\sim}{Z})$ contains all relevant information about the parameter (vector); and it will be seen that $N(\underset{\sim}{Z})$ contains all relevant information about the structure of the process. From $\delta(\underset{\sim}{Z}) = [S(\underset{\sim}{Z}), N(\underset{\sim}{Z})]$, one should almost always be able to reconstruct the original data, $\underset{\sim}{Z}$.

<u>Example 4.1</u>. (i) In Lemma 3.1, $S(\underset{\sim}{\xi}) = k\overline{\xi}$; $N(\underset{\sim}{\xi}) = \underset{\sim}{\xi}^*$, and $\delta(\underset{\sim}{\xi}) = [k\overline{\xi}, \underset{\sim}{\xi}^*]$ is the BDT. From $\delta(\underset{\sim}{\xi})$, one can reconstruct $\underset{\sim}{\xi}$.
(ii) In Lemma 3.2, $S(\underset{\sim}{\eta}) = \eta_k^*$ and $N(\underset{\sim}{\eta}) = \underset{\sim}{\eta}^{**}$. Hence $\delta(\underset{\sim}{\eta}) = \{\eta_k^*, \underset{\sim}{\eta}^{**}\}$.

From the examples above, it should be clear that there are several possible versions of the M-S-S and M-S-N.

<u>Remark</u>. The importance of these concepts is that as a rule of thumb in a goodness-of-fit problem, the decision rule should involve only the M-S-S while in a class-fit problem, the decision rule should only involve M-S-N. This will be seen in the sequel.

(B) <u>Distribution-free-ness</u>.

There are two types of distribution free statistics that arise in detection procedures.

Definition 4.2. (i) A statistic $T(\underset{\sim}{Z})$ is called nonparametric distribution free (NPDF) with respect to a family Ω^* of stochastic process laws if there exists a single distribution $Q(\cdot)$ such that

$$P\{T(\underset{\sim}{Z}) \leq t \mid L\} = Q(t) \quad \text{for all} \quad L \in \Omega^*,$$

(ii) A family of statistics, $\{T_1(\underset{\sim}{Z};\ L)\}$ indexed by the members $L \in \Omega^*$ is called parametric distribution free (PDF) with respect to Ω^* if there exists a distribution function $Q_1(\cdot)$ such that,

$$P\{T_1(\underset{\sim}{Z};\ L) \leq t\} = Q_1(t) \quad \text{for all} \quad L \in \Omega^*.$$

It is clear that each NPDF statistic is PDF.

Example 4.2. Let $\Omega^* = \Omega(\mathrm{PRP})$, and $\underset{\sim}{Z} = (Z_1, \ldots, Z_N)$ be the first N interarrival times with $Z_j = X_j$ for $1 \leq j \leq m$, and $Z_{m+r} = Y_r$ for $1 \leq r \leq n$, where $N = m + n$. Let $\hat{s}_1$ and $\hat{s}_2$ be the MLE's given by Theorem 2.2; $T_1 = \dfrac{m\hat{s}_2}{n\hat{s}_1}$ and $T_2 = \dfrac{2ms}{\hat{s}_1}$. Then, T_1 is NPDF wrt Ω^* with $Q = F(2m, 2n)$; and T_2 is PDF wrt Ω^* with $Q_1 = \chi^2_{2m}$. Furthermore, it can be shown that T_1 is a function of the $\underset{\sim}{Z}$ only through the M-S-N, $N(\underset{\sim}{Z})$; and T_2 is a function of $\underset{\sim}{X} = (X_1, \ldots, X_m)$ only through the M-S-S, $S(\underset{\sim}{X})$.

The M-S-N for the respective cases are given in Table 2.1. The relations between the M-S-S, M-S-N and DF statistics is best given by the following theorem.

Theorem 4.1. Let Ω^* be a family of cdfs admitting a M-S-S, $S(\underset{\sim}{Z})$, for data $\underset{\sim}{Z} = (Z_1, \ldots, Z_n)$. Then, (i) $T(\psi, \underset{\sim}{Z}) = \psi[N(\underset{\sim}{Z})]$ is NPDF wrt Ω^* for each (measurable) function $\psi(\cdot)$; and

(ii) $T^*(\psi^*, G, S(\underset{\sim}{Z})) = \psi^*[G, S(\underset{\sim}{Z})]$, when $\underset{\sim}{Z}$ is governed by G, is PDF wrt Ω^* for each (measurable) function $\psi^*(\cdot)$.

5. The K-S (Kolmogorov-Smirnov), Lilliefors and Srinivasan Statistics

(A) Kolmogorov's Original Statistic

Kolmogorov (1933) introduced the K-S statistic $D_n(F_0) = \sup_z |F_n(z) - F_0(z)|$, for continuous cdfs $F_0(\cdot)$ and empirical cdfs $F_n(\cdot)$, where $F_n(z) = \frac{1}{n} \sum_{j=1}^{n} \varepsilon\,(z - X_j) = \frac{1}{n} \sum_{j=1}^{n} \varepsilon(z - X(j))$, and $\varepsilon(u) = 1$ if $u \geq 0$; and $= 0$ if $u < 0$.

Definition 5.1. If $X_1, \ldots, X_n$ are i.i.d. F_0, continuous, then $D_n(F_0) \sim$ K-S(n).

In order to apply the K-S statistic directly, one must know $F_0(\cdot)$ completely. However, in many signal detection problems, $F_0(\cdot)$ is known only up to a nuisance parameter, or, equivalently is known only to be a member of a specific (parametric) family. Lilliefors (1967, 1969), Srinivasan (1970), and Choi (1980) introduced modified versions of the K-S statistic for such situations.

(B) Lilliefors-type Statistics

Let $\Omega^* = \{F(\theta;\cdot): \theta \in \Theta\}$ be a family of cdfs admitting a MLE, $\hat{\theta} = \hat{\theta}(X_1, \ldots, X_n)$, for θ.

Definition 5.2. (i) $\hat{F}(\cdot)$ is the cdf satisfying $\hat{F}(z) = F(\hat{\theta};z)$ for all z; and

(ii) $\hat{D}_n = \sup_z |F_n(z) - \hat{F}(z)|$.

Lilliefors (1967, 1969) calculated Monte Carlo tables for $\hat{D}_n$ in the normal and exponential cases, while Y. Choi (1980) has given such a table in the uniform case.

Srinivasan (1970) replaces $\hat{F}(\cdot)$ with the Rao-Blackwell estimate of $F(\theta;\cdot)$ in the Lilliefors statistic.

(C) Srinivasan-type Statistics

Consider a family $\Omega^* = \{F(\theta;\cdot): \theta \in \Theta\}$ of cdfs admitting a M-S-S, $S(\underset{\sim}{Z})$ for θ.

Definition 5.3. (i) $\tilde{F}(z) = P\{X_1 \le z | S(\underset{\sim}{Z})\}$; and

(ii) $\tilde{D}_n = \sup_z |F_n(z) - \tilde{F}(z)|$.

Srinivasan (1970) computed critical values of $\tilde{D}_n$ by (Monte Carlo) simulation for the exponential and normal families. Some of his numerical results however were in error, as was pointed out by Schafer, Finkelstein and Collins (1972).

Remark. These three statistics are in many cases, asymptotically equivalent. Calculations of $\tilde{F}_n(z)$ for the case when A is unknown, $s = s_0$ known and the case when both A and s are unknown are presented in Appendix A. These statistics are summarized in Table 5.1.

TABLE 5.1. Kolmogorov-Smirnov, Lilliefors, Srinivasan Statistics for Pareto Renewal Processes

Let $X_1, X_2, \ldots, X_n$ be i.i.d. Pa(A,s).

	K-S Statistic with M-S-N: $D_n^{(i)}$	Lilliefors Statistic: $\hat{D}_n$	Srinivasan Statistic: $\tilde{D}_n$
A unknown $s=s_0$ known	$D_n^{(i)} = \sup_z \left\lvert \frac{1}{n-1} \sum_{j=1}^{n-1} \varepsilon\,(z-Q_j^{(i)}) - F_j(z) \right\rvert$ $i = 1,2,3,4$ where $Q_j^{(1)} = V_j$, $F_1(z)=1-e^{-s_0 z}$ $Q_j^{(2)} = U_j$, $F_2(z)=1-z^{-s_0}$ $Q_j^{(3)} = W_j$, $F_3(z)=z$ $Q_j^{(4)} = Y_j$, $F_4(z)=z$ $\{V_j\}, \{U_j\}, \{W_j\}, \{Y_j\}$ defined in definition 3.2.	$\hat{D}_n = \sup_z \left\lvert \frac{1}{n} \sum_{j=1}^{n} \varepsilon\,(z-X(j)) - \hat{F}_n(z) \right\rvert$ where $\hat{F}_n(z) = 1 - \left(\frac{X(1)}{z}\right)^{s_0}$ for $z>X(1)$	$\tilde{D}_n = \sup_z \left\lvert \frac{1}{n} \sum_{j=1}^{n} \varepsilon\,(z-X(j)) - \tilde{F}_n(z) \right\rvert$ where $\tilde{F}_n(z) = 1-\left(\frac{n-1}{n}\right)\left(\frac{X(1)}{z}\right)^{s_0}$ for $z>X(1)$
$A=A_0$ known s unknown	$D_n = \sup_z \left\lvert \frac{1}{n-1} \sum_{j=1}^{n-1} \varepsilon\left(z-\frac{T_j}{T_n}\right) - z \right\rvert$ where $\{T_j\}$ is defined in Definition 3.2.	$\hat{D}_n$ same as above with $\hat{F}_n(z)=1-\left(\frac{A_0}{z}\right)^{\delta}$ for $z > A_0$, where $\delta = n\left[\sum_{j=1}^{n} \ln X_j - n \ln A_0\right]^{-1}$	$\tilde{D}_n$ same as above with $\tilde{F}_n(z)=1-(1-\delta')^{n-1}$ for $z > A_0$, where $\delta' = \frac{\ln z - \ln A_0}{\sum_{i=1}^{n} \ln X_i - n \ln A_0}$

TABLE 5.1. (Continued)

A,s both unknown	$D_n^{(1)} = \sup_z \left\lvert \frac{1}{n-2} \sum_{j=1}^{n-2} \varepsilon\left(z - \frac{\mathcal{D}_j}{\mathcal{D}_{n-1}}\right) - z \right\rvert$ $D_n^{(2)} = \sup_z \left\lvert \frac{1}{n-2} \sum_{j=1}^{n-2} \varepsilon\left(z - \frac{E_j}{E_{n-1}}\right) - z \right\rvert$ where $\{\mathcal{D}_j\}$, $\{E_j\}$ are defined in Definition 3.2.	$\hat{D}_n$ same as above with $\hat{F}_n(z) = 1 - \left(\frac{X(1)}{z}\right)^{\delta}$ for $z > X(1)$ where $\delta = n\left[\sum_{j=1}^{n} \ln X_j - n \ln X(1)\right]^{-1}$	$\tilde{D}_n$ same as above with $\tilde{F}_n(z) = \begin{cases} 0 & z < X(1) \\ 1/n & z = X(1) \\ 1 - \left(\frac{n-1}{n}\right)\left[1 - \frac{\delta}{n} \ln \frac{z}{X(1)}\right]^{n-1} & X(1) < z < X(1)e^{n/\delta} \\ 1 & z \geq X(1)e^{n/\delta} \end{cases}$ where $\delta = n\left[\sum_{i=1}^{n} \ln X_i - n \ln X(1)\right]^{-1}$

One can now treat the detection problems.

6. One-Sample Detection Procedures.

In this section, one will derive detectors for deciding between PN and the alternative N + S (noise plus signal). The data set will be denoted by $\underset{\sim}{Z} = (X_1, X_2, \ldots, X_n)$ and α is the PFA. There are altogether 8 problems considered. The first 4 deal with the cases when at least one of A and s is known and the last 4 concern cases in which both A and s are unknown. Numerical examples for each case are provided in the Appendix B, where they are numbered the same as the cases they illustrate. The results in this section are summarized in Table 6.1 at the end of this section.

Case 6.1. (A unknown, $s = s_0$ known, $L(A,s_0) \in \Omega(PRP)$)

$$\underline{PN}: A = A_0 \qquad \text{vs.} \qquad \underline{N+S}: A \neq A_0$$

The minimum PDF procedure is to decide N + S if and only if $X(1) < A_0$ or $X(1) > bA_0$ where $b = \alpha^{-1/ns_0}$. Equivalently, one can use the statistic

$$T = 2ns_0 \ln \frac{X(1)}{A_0}$$

which has a χ_2^2 distribution under PN. Note that the procedure is based solely on the M-S-S for A, which is $\hat{A} = X(1)$. The one-sided detection procedure for PN: $A \leq A_0$ vs. N + S: $A > A_0$ may be formulated similarly.

<u>Case 6.2</u>. (A unknown, $s = s_0$ known; class-fit problem)

$$\underline{\text{PN}}: \quad L(A,s_0) \in \Omega(\text{PRP}) \quad \text{vs.} \quad \underline{\text{N + S}}: \quad L(A,s_0) \notin \Omega(\text{PRP}).$$

One employs the K-S statistic with size $n - 1$ (K-S($n - 1$)) through anyone of the four versions of M-S-N, $\{V_j\}$, $\{U_j\}$, $\{W_j\}$, $\{Y_j\}$ given in Theorem 3.5 and Definition 3.3. Explicitely, they are

$$D_n^{(1)} = \sup_{z>0} \left| \frac{1}{n-1} \sum_{j=1}^{n-1} \varepsilon(z - V_j) - (1 - e^{-s_0 z}) \right|$$

$$D_n^{(2)} = \sup_{z>1} \left| \frac{1}{n-1} \sum_{j=1}^{n-1} \varepsilon(z - U_j) - (1 - z^{-s_0}) \right|$$

$$D_n^{(3)} = \sup_{0<z<1} \left| \frac{1}{n-1} \sum_{j=1}^{n-1} \varepsilon(z - W_j) - z \right|$$

$$D_n^{(4)} = \sup_{0<z<1} \left| \frac{1}{n-1} \sum_{j=1}^{n-1} \varepsilon(z - Y_j) - z \right|$$

where $D_n^{(i)} \sim$ K-S($n - 1$) for $i = 1, 2, 3, 4$. Therefore one decides N + S if and only if $D_n^{(i)} > d'$ where d' is an appropriate critical value from the K-S($n - 1$) distribution.

Alternatively, one may use the Srinivasan-type statistic (see Table as follows.

$$\tilde{D}_n = \left(\frac{n-1}{n}\right) \sup_{1<u<\infty} \left|\frac{1}{n-1} \sum_{j=1}^{n-1} \varepsilon\,(u - U_j) - (1 - u^{-s_0})\right|.$$

Note that $\tilde{D}_n \overset{d}{=} \left(\frac{n-1}{n}\right) D_n^{(2)}$.

The Lilliefors-type statistic is asymptotically equivalent to the Srinivasan-type statistic in this case.

Case 6.3. ($A = A_0$ known, s unknown, $L(A_0,s) \in \Omega(PRP)$)

PN: $s = s_0$ vs. N + S: $s \neq s_0$

The M-S-S for s is $\hat{s}=n\left[\sum_{j=1}^{n} \ln \frac{X(j)}{A_0}\right]^{-1}$ and the statistic $X(j)$ is $T = 2s_0 \sum_{j=1}^{n} \ln \frac{X(j)}{A_0} \sim \chi^2_{2n}$ under PN. Thus the detection procedure is decide N + S if and only if $T > \chi^2_{(2n,1-\alpha/2)}$ or $T < \chi^2_{(2n,\alpha/2)}$. Note that T depends on $\underset{\sim}{Z}$ only through $\hat{s}$. The one-sided procedure for PN: $s \leq s_0$ vs. N + S: $s > s_0$ may be formulated similarly.

Case 6.4. ($A = A_0$ known, s unknown; class-fit problem)

PN: $L(A_0,s) \in \Omega(PRP)$ vs. N + S: $L(A_0,s) \notin \Omega(PRP)$.

Again we should use the K-S statistic with the M-S-N which according to Table 2.1 is $\left\{\frac{T_1}{T_n}, \frac{T_2}{T_n}, \ldots, \frac{T_{n-1}}{T_n}\right\} \sim$ U-O-S$(n - 1)$.

Thus let $D_n = \sup_{0<z<1} \left|\frac{1}{n-1} \sum_{j=1}^{n-1} \varepsilon\left(z - \frac{T_j}{T_n}\right) - z\right|$ and one decides

$N \to S$ if and only if $D_n > d'$ where d' is the appropriate critical value from the K-S(n - 1) table.

<u>Case 6.5</u>. (A, s both unknown, $L(A,s) \in \Omega(PRP)$; Goodness-of-fit test)

$$\underline{PN}: \ L(A,s) = L(A_0,s_0) \quad \text{vs.} \quad \underline{N \to S}: \ L(A,s) \neq L(A_0,s_0)$$

From Theorem 3.3 (vi) and under PN situation,

$$\left\{\left[\frac{A_0}{X(n)}\right]^{s_0}, \ \left[\frac{A_0}{X(n-1)}\right]^{s_0}, \ \ldots, \ \left[\frac{A_0}{X(1)}\right]^{s_0}\right\} \sim \text{U-O-S}(n).$$

The statistic is (as in Case 6.3).

$$T = -2 \sum_{j=1}^{n} \ln\left[\frac{A_0}{X(j)}\right]^{s_0} \sim \chi^2_{2n} \quad \text{and the decision}$$

rule is decide $N \to S$ if and only if $T > \chi^2_{(2n,1-\alpha/2)}$ or $T < \chi^2_{(2n,\alpha/2)}$.

<u>Case 6.6</u>. (A and s unknown, $L(A,s) \in \Omega(PRP)$)

$$\underline{PN}: \ s \leq s_0 \quad \text{vs.} \quad \underline{N \to S}: \ s > s_0$$

The detection statistic here should only involve the M-S-S $S(\underset{\sim}{Z}) = \hat{s} = n\left[\sum_{j=1}^{n} \ln \frac{X(j)}{X(1)}\right]^{-1}$ for s.

The decision rule is: Decide $N \to S$ iff $\dfrac{2s_0}{\hat{s}} < c'$, where c' is the (100α)th percentile of the $\chi^2_{(2n-2)}$-distribution.

One notes that if $Y_j = \ln X_j$, then the Y's are i.i.d. T-exp (ln A,s). Park and U. Choi (1978) derive the minimum PFD one-sided procedures for the shape parameter s. It will be shown in Appendix A.3, that the Park-Choi decision rule is equivalent to the one given above.

For the two-sided detection problem: PN: $s = s_0$ vs. N + S: $s \neq s_0$, the decision rule follows from that above.

Case 6.7. (A, s unknown; $L(A,s) \in \Omega(PRP)$)

$$\underline{PN}: \; A \leq A_0 \quad \text{vs.} \quad \underline{N+S}: \; A > A_0$$

One should employ the M-S-S, $(\hat{A}, \hat{s})$, and the statistic

$$T(\underset{\sim}{Z}, A) = \frac{n(n-1)[\ln X(1) - \ln A]}{\sum_{j=1}^{n} \ln X_j - n \ln X(1)} \sim F(2, 2n-2).$$

The decision rule is: Decide N + S iff $X(1) > A_0$ and $T(\underset{\sim}{Z}, A_0) > f'$, where f' is the $100(1-\alpha)$th percentile of the $F(2, 2n-2)$-distribution.

Again, as in Case 6.6, Park and U. Choi (1978) give a minimum PDF procedure for the truncated exponential case, which is equivalent to the decision rule above.

Case 6.8. (A, s unknown; class-fit problem)

$$\underline{PN}: \; L(A,s) \in \Omega(PRP) \quad \text{vs.} \quad \underline{N+S}: \; L(A,s) \notin \Omega(PRP).$$

One can use the Kolmogorov-Smirnov, Lilliefors or Srinivasan-type statistics. From Theorem 3.4, one finds

$$D_n^{(1)} = \sup_{0<z<1} \left| \frac{1}{n-2} \sum_{j=1}^{n-2} \varepsilon\left(z - \frac{\mathcal{D}_j}{\mathcal{D}_{n-1}}\right) - z \right|$$

$$D_n^{(2)} = \sup_{0<z<1} \left| \frac{1}{n-2} \sum_{j=1}^{n-2} \varepsilon\left(z - \frac{E_j}{E_{n-1}}\right) - z \right|,$$

then $D_n^{(i)} \sim$ K-S(n - 2) for $i = 1, 2$. The decision rule is decide N + S if and only if $D_n^{(i)} > d'$ where d' is the critical value from the K-S(n - 2) table.

One now considers the Srinivasan statistic. From Table 5.1, one has

$$\tilde{D}_n = \sup_z |F_n(z) - \tilde{F}_n(z)|$$

$$= \sup_{a < z < ae^{n/s'}} \left|\frac{1}{n} \sum_{j=1}^{n} \varepsilon\left(\frac{z}{X(1)} - \frac{X(j)}{X(1)}\right) - 1 + \left(\frac{n-1}{n}\right)\left[1 - \frac{s'}{n} \ln \frac{z}{a}\right]^{n-1}\right|$$

where $\hat{A} = X(1) = a$, $\hat{s} = s'$. Let $v = \frac{z}{X(1)}$, $V(j) = \frac{X(j+1)}{X(1)}$. Then

$$\tilde{D}_n = \sup_{1 < v < e^{n/s'}} \left|\frac{1}{n} + \frac{1}{n}\left[\sum_{j=1}^{n-1} \varepsilon(v - V(j))\right] - 1 + \left(\frac{n-1}{n}\right)\left[1 - \frac{s'}{n} \ln v\right]^{n-1}\right|$$

$$= \sup_{1 < v < e^{n/s'}} \left|\frac{1}{n}\left[\sum_{j=1}^{n-1} \varepsilon(\ln v - \ln V(j))\right] - \left(\frac{n-1}{n}\right) + \left(\frac{n-1}{n}\right)\left[1 - \frac{s'}{n} \ln v\right]^{n-1}\right|.$$

Let $u = \ln v$, $U(j) = \ln V(j)$, then since $s = \dfrac{n}{\sum_{j=1}^{n-1} \ln V_j} = \dfrac{n}{(n-1)\bar{u}}$,

one has $\tilde{D}_n = \left(\frac{n-1}{n}\right) \sup_{0 < u < n/s'} \left|\frac{1}{n-1} \sum_{j=1}^{n-1} \varepsilon(u - U(j)) - \{1 - [1 - \frac{s'}{n} u]^{n-1}\}\right|$

$$= \left(\frac{n-1}{n}\right) \sup_{0 < u < (n-1)\bar{u}} \left| \frac{1}{n-1} \sum_{j=1}^{n-1} \varepsilon\,(u - U(j)) - \{1 - [1 - \frac{u}{(n-1)\bar{u}}]^{n-1}\} \right|$$

$$= \left(\frac{n-1}{n}\right) \sup_{0 < u < \sum_{j=1}^{n-1} U_j} \left| \frac{1}{n-1} \sum_{j=1}^{n-1} \varepsilon\,(u - U(j)) - \{1 - [1 - \frac{u}{(n-1)\bar{u}}]^{n-1}\} \right|.$$

The critical values of this statistic cannot be obtained from the known existing tables. They may be obtained by the Monte Carlo simulation method for various sample sizes and PFA .

These statistics are summarized in Table 6.1.

7. Two-Samples Detection Procedures

The data set here is $\underset{\sim}{Z} = (\underset{\sim}{X}, \underset{\sim}{Y}) = (X_1, \ldots, X_m, Y_1, \ldots, Y_n) = (Z_1, Z_2, \ldots, Z_N)$ where $N = m + n$. Here $\underset{\sim}{X}, \underset{\sim}{Y}$ are two independent random samples from $Pa(A_1, s_1)$ and $Pa(A_2, s_2)$ respectively. The letter α will denote the PFA. As in the one-sample case, numerical examples are given for each case in Appendix B with the same order and numberings as they are presented here.

Case 7.1. (A_1, A_2 known, s_1, s_2 unknown)

PN: $s_1 = s_2$ vs. N + S: $s_1 \neq s_2$

The M-S-S for (s_1, s_2) is $(\hat{s}_1, \hat{s}_2) = \left(m\left[\sum_{j=1}^{m} \ln\left(\frac{X_j}{A_1}\right)\right]^{-1}, n\left[\sum_{j=1}^{n} \ln\left(\frac{Y_j}{A_2}\right)\right]^{-1}\right)$

TABLE 6.1. **One-Sample Detection Procedures** (Data: $X_1, \ldots, X_m$ i.i.d. $Pa(A,s)$)

Case	Assumptions	PN	N + S	Statistics; PN-distribution	Decision Rule: Decide N + S iff
6.1	A unknown $s=s_0$ known $L(A,s_0)\varepsilon\Omega(PRP)$	$A = A_0$	$A \neq A_0$	(i) $X(1)$ or (ii) $T=2ns_0 \ln\frac{X(1)}{A_0} \sim \chi^2_2$	(i) $X(1)<A_0$ or $X(1)>bA_0$, $b = \alpha^{-1/ns_0}$ (ii) $T < 0$ or $T > \chi^2_{2,(1-\alpha)}$
6.2	A unknown $s=s_0$ known	$L(A,s_0)\varepsilon\Omega(PRP)$	$L(A,s_0)\notin\Omega(PRP)$	The K-S statistic $D_n^{(i)}$, i=1,2,3,4 are defined in Table 5.1	$D_n^{(i)} > d^*$, i=1,2,3,4, d^* appropriate critical value from K-S(n-1) table.
6.3	$A=A_0$ known s unknown $L(A_0,s)\varepsilon\Omega(PRP)$	$s = s_0$	$s \neq s_0$	$T=2s_0\sum_{j=1}^{n} \ln\frac{X(j)}{A_0} \sim \chi^2_{2n}$	$T > \chi^2_{2n,1-\alpha/2}$ or $< \chi^2_{2n,\alpha/2}$
6.4	$A=A_0$ known s unknown	$L(A_0,s)\varepsilon\Omega(PRP)$	$L(A_0,s)\notin\Omega(PRP)$	$D_n=\sup_{0<z<1}\left\vert\frac{1}{n-1}\sum_{j=1}^{n-1}\varepsilon(z-\frac{j}{n})-z\right\vert$ $\sim$ K-S(n-1)	$D_n > d'$ where d' is an appropriate critical value from the K-S(n-1) table.
6.5	A unknown s unknown $L(A,s)\varepsilon\Omega(PRP)$	$L(A,s)=L(A_0,s_0)$	$L(A,s)\neq L(A_0,s_0)$	$T=-2\sum_{j=1}^{n}\ln\left[\frac{A_0}{X(j)}\right]^{s_0} \sim \chi^2_{2n}$	$T > \chi^2_{2n,1-\alpha/2}$ or $< \chi^2_{2n,\alpha/2}$
6.6	A unknown s unknown $L(A,s)\varepsilon\Omega(PRP)$	$s \leq s_0$	$s > s_0$	$T=2s_0\{\sum_{j=1}^{n} \ln X_j - n\ln X(1)\}$	$T < \chi^2_{2n-2,\alpha}$

TABLE 6.1. (Continued)

6.7	A unknown s unknown $L(A,s)\varepsilon\Omega(PRP)$	$A \leq A_0$	$A > A_0$	$T = \dfrac{n(n-1)[\ln X(1) - \ln A_0]}{\sum_1^n \ln X_j - n\ln X(1)}$	$X(1) > A_0$ and $T > f_{(2,2n-2;1-\alpha)}$
6.8	A,s unknown	$L(A,s)\varepsilon\Omega(PRP)$	$L(A,s)\notin\Omega(PRP)$	$D_n^{(1)} =$ $\sup_{0<z<1} \left\| \frac{1}{n-2} \sum_{j=1}^{n-2} \varepsilon(z - \frac{\mathcal{D}_j}{\mathcal{D}_{n-1}}) - z \right\|$ $D_n^{(2)} =$ $\sup_{0<z<1} \left\| \frac{1}{n-2} \sum_{j=1}^{n-2} \varepsilon(z - \frac{E_j}{E_{n-1}}) - z \right\|$	$D_n^{(i)} > d^*$ i=1 or 2, d^* appropriate critical value from K-S(n-2) table.

and under PN, $T^* = \dfrac{n \sum_{j=1}^{m} \ln \frac{X_j}{A_1}}{m \sum_{j=1}^{n} \ln \frac{Y_j}{A_2}} \sim F(2m, 2n)$. Therefore,

decide $N + S$ if and only if $T > f'$ or $T < f''$ where f' and f'' are the appropriate percentiles of the $F(2m, 2n)$-distribution.

An important point arises in Case 7.1 above. The detection statistic T^* involved the M-S-S $(\hat{s}_1, \hat{s}_2)$ for (s_1, s_2), the unknown parameter pair. However, the remark following Example 4.1, and the cases of Section 6, suggest that for Case 7.1, one should employ the M-S-N, since the particular values of s_1, and s_2 are not pertinent here. This, indeed, is the case, as will be seen from the derivation to follow.

One can directly verify (from Lemma 3.1)

<u>Theorem 7.1</u>. Let $W_j = \ln \frac{Z_j}{A_1}$ for $1 \leq j \leq m$; and $= \ln \frac{Z_j}{A_2}$ for $m + 1 \leq j \leq N$; and $V_r = [\sum_1^r W_j] [\sum_1^N W_j]^{-1}$ for $1 \leq r \leq N - 1$. Then

(i) $(W_1, \ldots, W_N)$ are i.i.d. Exp(s).

(ii) $(V_1, \ldots, V_{N-1})$ is the M-S-N; and is $\sim$ U-O-S$(N - 1)$. Further,

(iii) $T^* = (\frac{n}{m}) \cdot (\frac{V_m}{1-V_m})$.

This means that T^* is a function of the data, $\underset{\sim}{Z}$, solely via the M-S-N of the combined sample. Hence, T^* is both a function of the

M-S-S's for the individual samples and the M-S-N of the combined samples.

The one-sided detection problems may be handled analogously.

Case 7.2. (A_1, A_2 unknown, s_1, s_2 known)

$$\underline{PN}: \quad A_1 = A_2 \qquad \text{vs.} \qquad \underline{N + S}: \quad A_1 \neq A_2$$

In this case, the M-S-S for (A_1, A_2) is $(\hat{A}_1, \hat{A}_2) = (X(1), Y(1))$. Since $\ln(\frac{X(1)}{A_1}) \sim \text{Exp}(ms_1)$, $\ln(\frac{Y(1)}{A_2}) \sim \text{Exp}(ns_2)$, one has by Lemma 3.6 the following,

Lemma 7.2. Let $T = \ln(\frac{Y(1)}{X(1)})$. Then under PN, the distribution of T is given by

$$H(t) = \begin{cases} 1 - \dfrac{ms_1}{(ms_1 + ns_2)} e^{-ns_2 t} & t \geq 0 \\[2ex] \dfrac{ns_2}{(ms_1 + ns_2)} e^{ms_1 t} & t < 0 \end{cases}$$

The proof of this lemma is given in Appendix A. Thus the decision rule is: Decide N + S if and only if $T > C_1$ or $T < C_2$ where C_1, C_2 are determined by $H(C_1) = 1 - \frac{\alpha}{2}$, and $H(C_2) = \frac{\alpha}{2}$.

In the special case where $s_1 = s_2 = s_0$ is known, the decision rule reduces to: Decide N + S iff $\frac{Y(1)}{X(1)} < b_1$ or $> b_2$ where $b_1 = [\frac{N}{n}(\frac{\alpha}{2})]^{\frac{1}{ms_0}}$ and $b_2 = [\frac{N}{m}(\frac{\alpha}{2})]^{\frac{1}{-ns_0}}$.

One further notes that $X^{-s} \sim U(0,\theta)$ with $\theta = A^{-s}$. For this uniform case, Y. Choi, Bell, Ahmad and Park (1982) present a detection procedure which coincides with the special case above.

<u>Case 7.3</u>. ($A_1 = A_2 = A$, unknown; s_1, s_2 unknown)

$$\underline{\text{PN}}: \quad s_1 = s_2 \qquad \text{vs.} \qquad \underline{\text{N + S}}: \quad s_1 \neq s_2$$

One first attempts to base the decision rule on the theorem below.

<u>Theorem 7.3</u>. Let $X_1, \ldots, X_n$, $Y_1, \ldots, Y_n$ be i.i.d. $P(A,s)$. Then (i) $X(1)$, $Y(1)$, $\eta_1 = \sum_{j=1}^{n-1} \ln \frac{X(j+1)}{X(1)}$ and $\eta_2 = \sum_{j=1}^{n-1} \ln \frac{Y(j+1)}{Y(1)}$

are independent; with

(ii) $2ns|\ln X(1) - \ln Y(1)| \sim \chi_2^2$; and

(iii) $2s\eta_1 \overset{d}{=} 2s\eta_2 \sim \chi_{2n-2}^2$

The decision rule for this case, with $m = n$, would then be:

Decide N + S iff

$$Q = \frac{2ns|\eta_1 - \eta_2|}{2ns|\ln X(1) - \ln Y(1)|} = \frac{|\eta_1 - \eta_2|}{|\ln X(1) - \ln Y(1)|} > C^*$$

However, the cdf of Q is not known, and, hence one seeks other approaches.

Beg (1980) derives a uniformly minimum PFD procedure for the case:

$$\underline{\text{PN}}: \quad s_1 \leq s_2 \qquad \text{vs.} \qquad \underline{\text{N + S}}: \quad s_1 > s_2$$

in the truncated-exponential case, which applies to the Pareto problem at hand even when $m \neq n$.

Let η_1 and η_2 be as above; $\eta = \sum_1^m \ln X_j + \sum_{j=1}^n \ln Y_j - NW$, where $N = n + m$ and $W = \min\{\ln X(1), \ln Y(1)\}$. Beg proves

Lemma 7.4. The conditional density of η_1, given W and η, is

$$h(\eta_1|w,\eta) = \frac{(m+n-2)!\eta_1^{m-2} (\eta - \eta_1)^{n-1}}{(m-2)!(n-1)!\ \eta^{m+n-2}} \quad \text{for } 0 < \eta_1 < \eta.$$

The decision rule based on this lemma becomes: Decide $N + S$ iff $\eta_1 < c = c(w,\eta)$ where $\int_c^\eta h(\eta_1|w, \eta)d\eta_1 = \alpha$.

If one performs the actual integration, it is easily seen that c satisfies the relation $c = c'z$ where $I_{c'}(m - 1, n) = 1 - \alpha$,

$$I_c(m,n) = \frac{1}{B(m,n)} \int_0^c y^{m-1}(1 - y)^{n-1}\, dy$$

is the incomplete Beta-function. A table for the incomplete Beta function has been tabulated by K. Pearson (1934).

Case 7.4. (A_1, A_2 unknown, $s_1 = s_2 = s$ unknown)

$$\underline{PN}: \quad A_1 \leq A_2 \qquad \text{vs.} \qquad \underline{N + S}: \quad A_1 > A_2$$

Let η_1, η_2, W, η be as defined in Case 7.3. Let $N = m + n$,

$$h(x^*|w,\eta) = \begin{cases} \frac{m}{N} & \text{if} \quad x^* = m \\ \frac{mn(N-2)}{N\eta}\left[1 - \frac{m(x^*-w)}{\eta}\right]^{N-3} & \text{if} \quad w < x^* < w + \frac{\eta}{m}, \end{cases}$$

and define the number $c = c(w, \eta)$ by

$$\int_{c}^{w+\frac{\eta}{m}} h(x^*|w,\eta)\,dx^* = \alpha.$$

From the result of Beg (1980), the decision rule is decide $N + S$ if and only if $\ln X(1) > c$. The number c may be found by performing the actual integration in which case one gets

$$c = \frac{\eta c'}{n} + w \quad \text{where} \quad c' = 1 - \left(\frac{N\alpha}{n}\right)^{\frac{1}{N-2}}$$

<u>Remark</u>: Suppose $m = n = \frac{N}{2}$, then Lemma 3.6 (vi) may be used to test <u>PN</u>: $s_1 = s_2$ vs. <u>N + S</u>: $s_1 \neq s_2$ in Case 7.3 and to test <u>PN</u>: $A_1 = A_2$ vs. <u>N + S</u>: $A_1 \neq A_2$ in Case 7.4. In both cases, under PN, the statistic

$$T = \frac{(N-2)N|\ln X(1) - \ln Y(1)|}{4(\eta_1 + \eta_2)} \sim F(2, 2N - 4).$$

The decision rule is decide $N + S$ if and only if $T > f'$ where f' is the appropriate percentile of the $F(2, 2N - 4)$-distribution.

<u>Case 7.5.</u> (A_1, A_2, s_1, s_2 all unknown)

<u>PN</u>: $s_1 = s_2$ vs. <u>N + S</u>: $s_1 \neq s_2$

Let $\eta_1 = \sum_{j=1}^{m-1} \ln \frac{X(j+1)}{X(1)}$, $\eta_2 = \sum_{j=1}^{n-1} \ln \frac{Y(j+1)}{Y(1)}$,

then from Lemma 3.6 (iii), under PN,

$$T = \frac{(n-1)\eta_1}{(m-1)\eta_2} \sim F(2(m-1),\ 2(n-1)).$$

The decision rule is decide N + S if and only if $T > f'$ or $T < f''$ where f' and f'' are appropriate percentiles of the $F(2m-2,\ 2n-2)$-distribution.

Case 7.6. (A_1, A_2, s_1, s_2 all unknown)

$$\underline{\text{PN}}:\quad A_1 = A_2 \qquad \text{vs.} \qquad \underline{\text{N + S}}:\quad A_1 \neq A_2$$

This detection procedure does not appear to have an elementary solution. One may try to apply the likelihood-ratio test. One has when not assuming $A_1 = A_2$,

$$\hat{A}_1 = X(1), \qquad \hat{s}_1 = m\left[\sum_{j=1}^{m-1} \ln \frac{X(j+1)}{X(1)}\right]^{-1}$$

$$\hat{A}_2 = Y(1), \qquad \hat{s}_2 = n\left[\sum_{j=1}^{n-1} \ln \frac{Y(j+1)}{Y(1)}\right]^{-1}.$$

Under PN, one has

$$A^* = \min\{X(1), Y(1)\} = Z(1) \quad \text{and}$$

$$s_1^* = m\left[\sum_{j=1}^{m-1} \ln \frac{X(j+1)}{Z(1)}\right]^{-1}, \qquad s_2^* = n\left[\sum_{j=1}^{n-1} \ln \frac{Y(j+1)}{Z(1)}\right]^{-1}.$$

Therefore the likelihood ratio is

$$\frac{\hat{L}_1}{\overset{*}{L}_0} = \begin{cases} \left(\frac{\hat{s}_2}{s_2^*}\right)^n \left(\frac{Y(1)}{X(1)}\right)^{\frac{\hat{s}_2 n}{s_2^*}} \left[\sum_{j=1}^{n} Y_j\right]^{(s_2^* - \hat{s}_2)} & \text{if } Z(1) = X(1) < Y(1) \\ \left(\frac{\hat{s}_1}{s_1^*}\right)^m \left(\frac{X(1)}{Y(1)}\right)^{\frac{\hat{s}_1 m}{s_1^*}} \left[\sum_{j=1}^{m} X_j\right]^{(s_1^* - \hat{s}_1)} . & \text{if } Z(1) = Y(1) < X(1) \end{cases}$$

The distribution of $\hat{L}_1/\hat{L}_0$ is complex and the above expression resembles in structure that of the Behrens-Fisher problem.

<u>Case 7.7</u>. (A_1, A_2, s_1, s_2 all *unknown*)

<u>PN</u>: $(A_1, s_1) = (A_2, s_2)$ vs. <u>N + S</u>: $(A_1, s_1) \neq (A_2, s_2)$

The decision rule consists of 2 steps, first deciding whether $s_1 = s_2$ and then if one decides $s_1 = s_2$ one tries to decide whether $A_1 = A_2$. The procedure is a combination of Cases 7.5 and 7.4.

Let $\eta_1 = \sum_{j=1}^{m-1} \ln \frac{X(j+1)}{X(1)}$, $\eta_2 = \sum_{j=1}^{n-1} \ln \frac{Y(j+1)}{Y(1)}$, then

under PN, $T = \frac{(n-1)\eta_1}{(m-1)\eta_2} \sim F(2(m-1), 2(n-1))$.

The decision rule is:

(i) Decide N + S if

$$T > f_{(2(m-1),2(n-1)1-\alpha/4)} \equiv C_1 \quad \text{or} \quad T < f_{(2(m-1),2(n-1),\alpha/4)} \equiv C_2$$

(ii) If $C_2 \le T \le C_1$, then decide $N + S$ if and only if $\ln X(1) > C_3$ and $\ln Y(1) > C_4$ where $C_3 = \frac{\eta C_3'}{n} + w$ $C_4 = \frac{\eta C_4'}{m} + w$,

$$w = \min\{\ln X(1), \ln Y(1)\}, \quad \eta = \sum_{j=1}^{m} \ln X_j + \sum_{j=1}^{n} \ln Y_j - Nw,$$

$$C_3' = 1 - \left(\frac{N\alpha}{4n}\right)^{\frac{1}{N-2}}, \quad C_4' = 1 - \left(\frac{N\alpha}{4m}\right)^{\frac{1}{N-2}}, \quad N = m + n.$$

Table 7.1 summarizes these procedures.

TABLE 7.1. Two-Samples Signal Detection Procedures

$\underset{\sim}{X} = (X_1, \ldots, X_m) \sim$ i.i.d. $Pa(A_1,s_1)$, $\underset{\sim}{Y} = (Y_1,\ldots, Y_n) \sim$ i.i.d. $Pa(A_2,s_2)$, $\underset{\sim}{X}$, $\underset{\sim}{Y}$ independent, $\eta_1 = \sum_{j=1}^{m} \ln X_j - m \ln X(1)$,

$\eta_2 = \sum_{j=1}^{n} \ln Y_j - n \ln Y(1)$, $N = m+n$, $\eta = \sum_{j=1}^{m} \ln X_j + \sum_{j=1}^{n} \ln Y_j - NW$, $W = \min\{\ln X(1), \ln Y(1)\}$

Cases	Assumptions	PN	N + S	Statistics; PN-distributions	Decision Rule-decide N+S iff
7.1	(A_1,A_2) known (s_1,s_2) unknown	$s_1 = s_2$	$s_1 \neq s_2$	$T = \dfrac{n\sum_{j=1}^{m} \ln(\frac{X_j}{A_1})}{m\sum_{j=1}^{n} \ln(\frac{Y_j}{A_2})} \sim F(2m,2n)$	$T^* > f_{(2m,2n,1-\alpha/2)}$ or $T^* < f_{(2m,2n,\alpha/2)}$
7.2	(A_1,A_2) unknown (s_1,s_2) known	$A_1 = A_2$	$A_1 \neq A_2$	$T=\ln(\frac{Y(1)}{X(1)})$, $T \sim H(x)$ where $H(x)=\begin{cases} 1-\dfrac{ms_1}{(ms_1+ns_2)} e^{-ns_2x} & x\geq 0 \\ \dfrac{ns_2}{(ms_1+ns_2)} e^{ms_1x} & x\leq 0 \end{cases}$	$T > C_1$ or $T < C_2$ where $H(C_1) = 1-\alpha/2$ $H(C_2) = \alpha/2$. If $s_1=s_2=s_0$, then test reduces to $Y(1)/X(1)<b_1$ or $> b_2$ where $b_1=[\frac{N}{n}(\frac{\alpha}{2})]^{1/ms_0}$, $b_2=[\frac{N}{m}(\frac{\alpha}{2})]^{-1/ns_0}$
7.3	$A_1=A_2=A$ unknown (s_1,s_2) unknown	$s_1 \leq s_2$	$s_1 > s_2$	C' determined by $I_{C'}(m-1,n) = 1 - \alpha$	$\eta_1 > C'\eta$
7.4	(A_1,A_2) unknown $s_1=s_2=s$ unknown	$A_1 \leq A_2$	$A_1 > A_2$	$c = \dfrac{\eta c'}{n} + w$ $c' = 1 - (\frac{N\alpha}{n})^{1/(N-2)}$	$\ln X(1) > c$

TABLE 7.1. (Continued)

7.5	(A_1, A_2, s_1, s_2) all unknown	$s_1 = s_2$	$s_1 \neq s_2$	$T = \dfrac{(n-2)\eta_1}{(m-2)\eta_2} \sim F(2(m-1), 2(n-1))$	$T > f_{(2(m-1),2(n-1),1-\alpha/2)}$ or $T < f_{(2(m-1),2(n-1),\alpha/2)}$
7.6	(A_1, A_2, s_1, s_2) all unknown	$A_1 = A_2$	$A_1 \neq A_2$	Unknown	Unknown
7.7	(A_1, A_2, s_1, s_2) all unknown	$(A_1, s_1) = (A_2, s_2)$	$(A_1, s_1) \neq (A_2, s_2)$	$c_1 = f_{(2(m-1),2(n-1),1-\alpha/4)}$ $c_2 = f_{(2(m-1),2(n-1),\alpha/4)}$ $c_3' = 1-(\frac{N\alpha}{4n})^{1/(N-2)}$ $c_4' = 1-(\frac{N\alpha}{4m})^{1/(N-2)}$ $c_3 = \frac{\eta c_3'}{n} + w$, $c_4 = \frac{\eta c_4'}{m} + w$	decide $N + S$ if $T > c_1$ or $T < c_2$. If $c_2 \leq T \leq c_1$, then decide $N + S$ if and only if $\ln X(1) > c_3$, and $\ln Y(1) > c_4$.

REFERENCES

1. Ahmad, R. and Abouammoh, A. M. (1977). "On the structure and applications of infinite divisibility, stability and symmetry in stochastic inference" Recent Developments in Statistics (eds. J. R. Barra et. al.), 303-317. North-Holland Publishing Co., Amsterdam.
2. Beg, M. A. (1981). Exponential family distribution with unknown truncation parameter: two sample problem. Manuscript, Monash University, Clayton, Australia.
3. Bell, C. B. (1964a). Automatic Distribution-Free Statistical Signal Detection. Research Report, U.S. Navy, Electronics Lab., San Diego.
4. Bell, C. B. (1964b). Some basic theorems for distribution-free statistics. Ann. Math. Statist; 35, 150-156.
5. Bergstrom, H. (1952). On some expansions of stable distributions, Arkiv fur Matematik, 2, 375-378.
6. Choi, Y., Bell, C. B., Ahmad, R. and Park, C. J. (1982). Signal detection for uniform renewal processes. Series in Statistics and Biostatistics #1-82, pp. 1-59, Department of Mathematical Sciences, San Diego State University.
7. Fama, E. F. and Roll, R. (1968). Some properties of symmetric stable distributions. J. Amer. Statist. Assoc., 63, 817-836.
8. Fisk, P. R. (1961). The graduation of income distributions, Econometrica, 29, 171-185.

9. Freiling, E. C. (1966). A comparison of the fallout mass-size distributions calculated by lognormal and power-law models. San Francisco: U.S. Naval Radiological Defense Laboratory.

10. Hagstroem, K. G. (1960). Remarks on Pareto distributions. Skandinavisk Aktuarietidskrift, 59-71.

11. Johnson, N. L. and Kotz, S. (1970). Continuous Univariate Distributions - I. Wiley.

12. Lehman, E. L. (1959). Testing Statistical Hypotheses. Wiley.

13. Levy, P. (1925). Calcul des Probabilities. Gauthier-Villas, Paris.

14. Lilliefors, H. W. (1967). On the Kolmogorov-Smirnov test for normality with mean and variance unknown. J. Amer. Statist. Assoc., 62, 399-402.

15. Lilliefors, H. W. (1969). On the Kolmogorov-Smirnov test for the exponential distribution with mean unknown. J. Amer. Statist. Assoc., 64, 387-389.

16. Malik, H. J. (1966). Exact moments of order statistics from the Pareto distribution. Skandinavisk Aktuarietidskrift, 144-157.

17. Malik, H. J. (1967). Exact moments of order statistics for a power-function population. Skandinavisk Aktuarietidskrift, 64-69.

18. Mandelbrot, B. (1960). The Pareto-Levy law and the distribution of income. International Economic Review, 1. 79-106.

19. Mandelbrot, B. (1963). The variation of some other speculative prices. J. of Business, University of Chicago, 36, 394-419

20. Muniruzzaman, A. N. M. (1950). On some distributions in connection with Pareto's law. Proc. 1st Pakistan Statistical Conf., 90-93.

21. Muniruzzaman, A. N. M. (1957). On measures of location and dispersion and tests of hypotheses on a Pareto population. Bull. Calcutta Statist. Assoc., 7, 115-123.

22. Pareto, Vilfredo (1897). Cours d'Economie Politique. Lausanne and Paris: Rouge and Cie.

23. Park, C. J. and Choi, U-Jin (1978). Minimum Variance unbiased estimat on for the negative exponential distribution with unknown truncation parameter. Proc. College of Natural Sciences, Seoul National University, Seoul, Korea; 3, 13-18.

24. Pearson, K. (1934). Tables of the Incomplete Beta Function. Cambridge University Press.

25. Pigou, A. C. (1932). The Economics of Welfare, Macmillan Company, London.

26. Pyke, R. (1965). Spacings (with Discussion). J. Roy. Statist. Soc. Ser. B, 27, 395-449.

27. Quandt, R. E. (1966). Old and new methods of estimation and the Pareto distribution. Metrika, 10, 55-82.

28. Renyi, A. (1953). On the theory of order statistics. Acta Math. Acad. Sci. Hung, 4, 191-231.

29. Rider, P. R. (1964). Distribution of product and of quotient of maximum values in samples from a power-function population. J. Amer. Statist. Assoc., 59, 877-880.

30. Schaffer, R. E., Finkelstein, J. M. and Collins, John (1972). "On a goodness-of-fit test for the exponential distribution with mean

unknown", Biometrika, 59, 222-224.

31. Srinivasan, R. (1970). An approach to testing the goodness-of-fit of incompletely specified distributions. Biometrika, 57, 605-611.

32. Steindl, J. (1965). Random Processes and Growth of Firms. Hafner, New York.

APPENDIX A: Proofs

A.1: A unknown, $s = s_0$ known, then
$$\tilde{F}_n(z) = \begin{cases} 1 - (\frac{n-1}{n})(\frac{X(1)}{z})^{s_0} & z > X(1) \\ 0 & z \le X(1). \end{cases}$$

Proof. $\tilde{F}_n(z) = E\{I_{\{X_1 \le z\}} | S(\underset{\sim}{Z})\} = P\{X_1 \le z | X(1)\}$

$$= \frac{1}{n} \sum_{j=1}^{n} P\{X_j \le z | X(1)\}$$

$$= \frac{1}{n} + \frac{1}{n} \sum_{j=1}^{n-1} P\{X(j+1) \le z | X(1)\} \qquad z > X(1)$$

$$= \frac{1}{n} + \frac{1}{n} \sum_{j=1}^{n-1} P\left\{\frac{X(j+1)}{X(1)} \le \frac{z}{X(1)}\right\}. \quad \text{from Theorem 3.5 (ii)}$$

Now let $U(j) = \frac{X(j+1)}{X(1)}$, then $\{U(1), U(2), \ldots, U(n-1)\} \sim Pa(1, s_0)$-O-S$(n - 1)$.

Therefore
$$\tilde{F}_n(z) = \frac{1}{n} + (\frac{n-1}{n})\left[\frac{1}{n-1} \sum_{j=1}^{n-1} P\{U_j \le \frac{z}{X(1)}\}\right]$$

$$= \frac{1}{n} + (\frac{n-1}{n})\left[1 - P\{U_1 > \frac{z}{X(1)}\}\right]$$

$$= 1 - (\frac{n-1}{n}) P\{U_1 > \frac{z}{X(1)}\}$$

$$= 1 - (\frac{n-1}{n})(\frac{X(1)}{z})^{s_0}.$$

A.2. A, s both unknown, then $\tilde{F}_n(z) = \begin{cases} 0 & z < a \\ \frac{1}{n} & z = a \\ 1 - (\frac{n-1}{n})[1 - \frac{s}{n}\ln\frac{z}{a}]^{n-1} & a < z < ae^{n/s} \\ 1 & z = ae^{n/s} \end{cases}$

where $\hat{A} = a$, $\hat{s} = s$.

Proof. $\tilde{F}_n(z) = E\{I_{\{X_1 \le z\}} | S(\tilde{Z})\} = P\{X_1 \le z | (\hat{A}, \hat{s})\}$

$$= \frac{1}{n} \sum_{j=1}^{n} P\{X_j \le z | (\hat{A}, \hat{s})\}$$

$$= \frac{1}{n} + \frac{1}{n} \sum_{j=1}^{n-1} P\{X(j+1) \le z | (\hat{A}, \hat{s})\}. \qquad z > a$$

Let $U(j) = \frac{X(j+1)}{X(1)}$, $V(j) = \ln U(j)$, then by Theorem 3.5,

$\{U(1), U(2), \ldots, U(n-1)\} \sim Pa(1, s)\text{-O-S}(n-1)$ and

$\{V(1), V(2), \ldots, V(n-1)\} \sim Exp(s)\text{-O-S}(n-1)$.

Hence $\tilde{F}_n(z) = \frac{1}{n} + \frac{1}{n} \sum_{j=1}^{n-1} P\{U(j) \le \frac{z}{a} | s\}$

$$= \frac{1}{n} + \frac{1}{n} \sum_{j=1}^{n-1} P\{V(j) \le \ln\frac{z}{a} | s\}$$

$$= \frac{1}{n} + \frac{1}{n} \sum_{j=1}^{n-1} P\{\frac{V_j}{(n-1)\overline{V}} \le \frac{s}{n} \ln\frac{z}{a}\}$$

since $s = \hat{s} = \dfrac{n}{\sum_{j=1}^{n-1} V_j} = \dfrac{n}{(n-1)\overline{V}}$.

By Lemma 3.1, $\{\frac{V_1}{(n-1)\overline{V}}, \frac{V_1 + V_2}{(n-1)\overline{V}}, \ldots\} \sim$ U-O-S$(n - 1)$. To continue,

$$\tilde{F}_n(z) = \frac{1}{n} + (\frac{n-1}{n}) P\{\frac{V_1}{(n-1)\overline{V}} \leq \frac{s}{n} \ln \frac{z}{a}\}$$

$$= 1 - (\frac{n-1}{n}) P \{\frac{V_1}{(n-1)\overline{V}} \geq \frac{s}{n} \ln \frac{z}{a}\}$$

$$= 1 - (\frac{n-1}{n})[1 - \frac{s}{n} \ln \frac{z}{a}]^{n-1} .$$

In the above derivation, since $V_j \leq (n - 1)\overline{V} = \frac{n}{s}$ for every j, one may restrict z to $\ln \frac{z}{a} \leq \frac{n}{s}$ or $z \leq ae^{n/s}$.

A.3. One verifies here that the u.m.p. procedure found by Park and U. Choi (1978) is equivalent to that in Cases 6.6 and 6.7. Park and Choi considered the p.d.f. of the truncated exponential,

$$f(x,\lambda,\nu) = \lambda e^{-\lambda(x-\nu)} I_{(\nu,\infty)}(x) \qquad 0 < \nu, \quad \lambda < \infty$$

where λ,ν are unknown. Let $Y_1, Y_2, \ldots, Y_n$ be random sample from a truncated exponential distribution. Let $S = \sum_{i=1}^{n} Y_i$, then given $Y(1) = y$,

$$f(s|y) = \frac{\lambda^{n-1}(s-ny)^{n-2} e^{-\lambda(s-ny)}}{\Gamma(n-1)} I_{(ny,\infty)}(s) \tag{1}$$

while given $S = s$,

$$f(y|s) = \frac{n(n-1)(s-ny)^{n-2}}{(s-n\nu)^{n-1}} I_{(\nu,s/n)}(y). \tag{2}$$

Theorem A. (Park and U. Choi). For testing hypothesis H_ν: $\nu \leq \nu$ against K_ν: $\nu > \nu_0$, the u.m.p. unbiased test is given by

$$\phi(Y(1)) = \begin{cases} 1 & \text{if } Y(1) \geq C(s) \\ 0 & \text{if } Y(1) < C(s) \end{cases}$$

where $C(s)$ is uniquely determined by $P\{Y(1) \geq C(s) | S = s, \nu = \nu_0\} = \alpha$.

Theorem B. (Park and U. Choi). For testing hypothesis H_λ: $\lambda \leq \lambda_0$ against, K_λ: $\lambda > \lambda_0$, the u.m.p. unbiased test is given by

$$\phi(S) = \begin{cases} 1 & \text{if } S \leq C(Y(1)) \\ 0 & \text{if } S > C(Y(1)) \end{cases}$$

where $C(Y(1))$ is uniquely determined by $P\{S \leq C(Y(1)) | Y(1) = y, \lambda = \lambda_0\} = \alpha$

From Theorem A and (2), one concludes $C = C(S)$ is determined by

$$\frac{n(n-1)}{(s-n\nu_0)^{n-1}} \int_C^{\frac{s}{n}} (s-ny)^{n-2}\, dy = (n-1)\int_{C'}^{1} (1-u)^{n-2}\, du$$

where $U = \dfrac{n(Y(1)-\nu_0)}{S-n\nu_0} \sim B(1, n-1)$. But

$$U = \frac{n(Y(1)-\nu_0)}{S-nY(1)+nY(1)-n\nu_0} = \frac{T'}{1+T'} \quad \text{where } T' = (n-1)T,$$

T is the statistic given in Section 6, Case 6.7. Since the function $f(t') = \frac{t'}{1 + t'}$ is strictly increasing for $t' > 0$, the two tests are equivalent. It is also easily verified that if $U \sim B(1, n - 1)$, then $\frac{T'}{n - 1} \sim F(2, 2(n - 1))$. The proof of the equivalence between Theorem B and the test procedure in Case 6.6 follows from (1) in the same manner and will be omitted.

A.4. (Proof of Lemma 7.1). From the hypothesis, let $U = \ln Y(1)$,

$V = \ln X(1)$ and $T = U - V$, then

$$
\begin{aligned}
P\{T \leq t\} &= P\{U - V \leq t\} \\
&= \int_0^{\infty} P\{U \leq v + t\}(ms_1)e^{-ms_1 v}\, dv \\
&= ms_1 \int_{\max(-t,0)}^{\infty} [1 - e^{-ns_2(v+t)}]e^{-ms_1 v}\, dv.
\end{aligned}
$$

Considering separately the case $t > 0$ and $t < 0$, the above integral can be evaluated and equals to $H(t)$ given in Lemma 7.2.

APPENDIX B: Numerical Examples

The numerical examples correspond to the various techniques developed in the main body of the report, and have corresponding numbers.

The data is divided into three sets.

Table 1. This consists of simulated interarrival times from Pa(1;1); Pa(2;1); Pa(3;1); and Pa(5;5) distributions.

Table 2. This consists of simulated interarrival times from Pa(1;2); Pa(1;5); Pa(2;2); and Pa(5;2) distributions.

Table 3. This consists of real data related to 24 complete heart-beat cycles. [Special attention will be paid to the waiting times for the "R-peaks". See sketch below the table.]

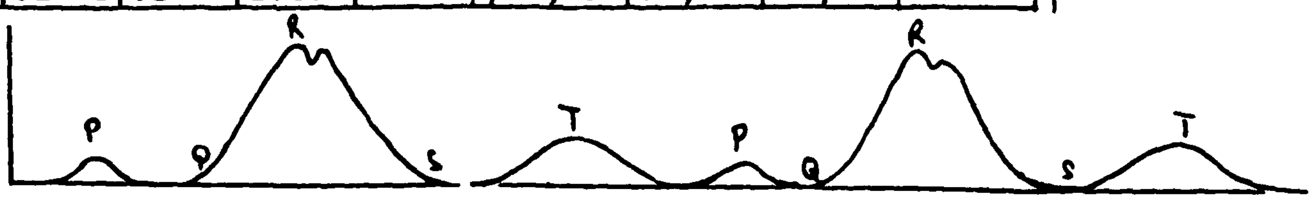
TABLE 1: Pareto Data (Interarrivals)

Pa(1;1)		Pa(2;1)		Pa(3;1)		Pa(5;5)	
1	1.724	1	3.146	1	9.867	1	5.019
2	1.052	2	2.416	2	4.718	2	11.065
3	1.272	3	6.496	3	3.701	3	6.738
4	3.360	4	2.886	4	38.658	4	5.221
5	2.204	5	3.828	5	5.777	5	5.405
6	159.695	6	2.192	6	9.975	6	5.355
7	3.627	7	2.570	7	8.182	7	6.681
8	3.272	8	6.280	8	3.711	8	5.537
9	1.451	9	4.440	9	21.597	9	5.006
10	2.613	10	3.401	10	5.968	10	5.285
11	7.524	11	6.176	11	11.192	11	5.398
12	1.202	12	4.808	12	19.364	12	6.091
13	1.715	13	3.523	13	11.310	13	12.695
14	10.139	14	47.815	14	21.817	14	5.817
15	3.616	15	24.820	15	6.251	15	6.879
16	1.612	16	3.798	16	3.397	16	5.325
17	11.976	17	3.060	17	9.255	17	5.560
18	1.010	18	3.074	18	39.874	18	5.533
19	1.021	19	2.108	19	5.310	19	6.811
20	1.441	20	3.285	20	9.901	20	5.754
21	1.070	21	2.123	21	3.984	21	5.125
22	4.715	22	2.207	22	3.441	22	9.944
23	7.654	23	3.194	23	3.290	23	7.005
24	1.159	24	2.744	24	18.728	24	7.560
25	1.221	25	6.041	25	9.260	25	12.671
26	1.849	26	2.291	26	6.073	26	7.046
27	53.193	27	5.100	27	3.921	27	5.387
28	3.184	28	9.542	28	4.276	28	6.155
29	1.307	29	3.426	29	10.197	29	6.094
30	1.062	30	4.741	30	4.103	30	5.921
31	3.457	31	2.350	31	12.132	31	5.416
32	1.709	32	6.036	32	3.008	32	7.273
33	1.107	33	6.720	33	16.134	33	6.041
34	1.128	34	3.535	34	3.140	34	6.104
35	2.013	35	5.530	35	10.051	35	5.032
36	2.522	36	5.950	36	3.155	36	6.131
37	1.147	37	2.128	37	22.246	37	5.491
38	1.074	38	3.654	38	3.501	38	6.569
39	1.679	39	3.201	39	3.338	39	5.752
40	4.771	40	5.268	40	6.149	40	11.101
41	1.126	41	3.608	41	41.302	41	5.465
42	2.155	42	7.745	42	3.149	42	10.545
43	1.592	43	7.062	43	3.094	43	5.388
44	2.239	44	10.085	44	70.014	44	5.208
45	2.184	45	4.212	45	3.448	45	5.025
46	4.257	46	2.033	46	4.124	46	5.034
47	1.310	47	3.831	47	76.268	47	7.443
48	1.691	48	61.096	48	7.460	48	5.045
49	2.253	49	6.118	49	3.001	49	9.862
50	2.020	50	4.532	50	9.305	50	6.499

TABLE 2 Pareto Data (Interarrivals)

Pa(1;2)		Pa(1;5)		Pa(2;2)		Pa(5;2)	
1	1.042	1	2.066	1	3.603	1	5.975
2	2.182	2	1.120	2	2.871	2	22.850
3	1.096	3	1.463	3	2.137	3	11.492
4	2.559	4	1.075	4	2.235	4	19.377
5	1.087	5	1.173	5	9.685	5	17.095
6	1.215	6	1.063	6	2.111	6	5.814
7	1.039	7	1.145	7	2.164	7	6.870
8	1.663	8	1.899	8	2.164	8	7.129
9	1.988	9	1.303	9	2.410	9	15.345
10	1.886	10	1.321	10	2.824	10	8.195
11	1.626	11	2.217	11	4.459	11	26.229
12	3.271	12	1.392	12	4.658	12	11.606
13	1.788	13	1.304	13	3.665	13	5.214
14	1.225	14	1.099	14	2.159	14	5.998
15	1.240	15	1.153	15	5.737	15	6.334
16	1.400	16	1.012	16	4.274	16	5.142
17	2.333	17	1.145	17	2.218	17	5.114
18	1.609	18	1.198	18	2.048	18	6.474
19	2.935	19	1.006	19	2.732	19	6.909
20	1.084	20	1.263	20	2.633	20	8.616
21	1.187	21	1.057	21	2.341	21	10.262
22	1.177	22	1.215	22	3.323	22	6.515
23	1.249	23	1.273	23	2.492	23	6.313
24	7.952	24	1.133	24	2.812	24	7.769
25	3.738	25	1.101	25	2.123	25	7.494
26	1.024	26	1.034	26	3.226	26	6.104
27	1.538	27	1.239	27	2.288	27	11.448
28	1.606	28	1.276	28	2.131	28	6.439
29	1.286	29	1.175	29	5.564	29	6.916
30	1.056	30	1.149	30	7.313	30	6.428
31	3.618	31	1.078	31	2.287	31	9.624
32	1.318	32	1.390	32	2.089	32	5.423
33	1.049	33	1.348	33	2.287	33	13.003
34	1.322	34	1.140	34	2.809	34	7.036
35	2.867	35	1.009	35	2.769	35	9.043
36	1.129	36	1.059	36	2.049	36	8.128
37	1.303	37	1.001	37	2.044	37	11.762
38	1.893	38	1.076	38	2.087	38	9.741
39	1.486	39	1.357	39	2.308	39	5.535
40	2.070	40	1.296	40	3.526	40	9.106
41	1.266	41	1.008	41	2.231	41	5.047
42	1.115	42	1.011	42	2.095	42	90.560
43	2.698	43	1.021	43	2.187	43	9.243
44	1.392	44	1.571	44	3.379	44	21.927
45	1.256	45	1.151	45	2.235	45	10.706
46	1.162	46	1.312	46	2.952	46	6.904
47	1.193	47	1.138	47	2.714	47	15.453
48	1.404	48	1.585	48	2.147	48	6.825
49	1.156	49	1.152	49	3.417	49	6.509
50	1.334	50	1.236	50	3.688	50	5.783

Cycle No	MAGNITUDE P	MAGNITUDE Q	MAGNITUDE R	MAGNITUDE T	TIMES P	TIMES Q	TIMES R	TIMES T	Cycle Length
1	.0927	.0708	1.190	.4785	148	428	467	743	1014
2	.1005	.0917	1.203	.4344	1,183	1,440	1,484	1,758	1014
3	.1121	.0982	1.212	.4316	1,897	2,378	2,498	2,884	1040
4	.1238	.0608	1.297	.5652	3,452	3,521	3,558	3,835	1041
5	.0819	.0206	1.261	.5060	4,108	4,559	4,599	4,812	1056
6	.0769	.0764	1.284	.4797	5,346	5,618	5,655	5,923	1078
7	.0815	.0507	1.228	.4700	6,346	6,689	6,733	7,005	1073
8	.1263	.1113	1.166	.4240	7,596	7,756	7,806	8,085	1076
9	.1009	.0817	1.274	.4542	8,645	8,840	8,882	9,154	1089
10	.0839	.0601	1.239	.4663	9,694	9,935	9,971	10,250	1044
11	.1126	.0907	1.164	.4482	10,727	10,985	11,015	11,301	1012
12	.0939	.0894	1.189	.4554	11,833	11,985	12,027	12,300	1028
	.0929	.0648	1.278	.4778	12,876	13,017	13,055	13,331	1038
14	.1010	.0555	1.213	.4742	13,758	14,050	14,093	14,367	1062
15	** .0560	.0374	1.235	.4874	14,307	15,112	15,153	15,429	1063
16	.0832	.0899	1.294	.4268	16,007	16,179	16,216	16,495	1083
17	.1278	.0939	1.294	.4465	17,053	17,265	17,299	17,587	1063
18	.1044	.0585	1.255	.5001	18,255	18,325	18,362	18,360	1018
19	.1184	.0702	1.250	.4491	19,286	19,345	19,380	19,655	1013
20	.1142	.1141	1.266	.4105	20,203	20,345	20,393	20,665	947
21	.1338	.1140	1.228	.4052	21,103	21,297	21,340	21,612	891
22	.1331	.1050	1.230	.4405	22,035	22,195	22,231	22,495	911
23	** .0849	.0584	1.127	.4265	23,068	23,097	23,140	23,415	901
24	.0916	.0807	1.156	.4038	23,825	24,002	24,041	24,308	947
25	.1442	.1323	1.260	--	24,766	24,937	24,986	--	



<u>Example 6.1</u>: A unknown, $s = s_0$ known, $L(A,s_0) \in \Omega(PRP)$

PN: $A = A_0$ vs. N + S: $A \neq A_0$.

<u>Detector Statistics</u>: (i) $X(1)$ or

(ii) $T = 2ns_0 \ln \frac{X(1)}{A_0} \sim \chi_2^2$.

<u>Decision Rule</u>: Decide N + S iff

(i) $X(1) < A_0$ or $X(1) > bA_0$, $b = \alpha^{-1/ns_0}$

(ii) $T < 0$ or $T > \chi^2_{2,(1-\alpha)}$.

Decision rule (i), the MP procedure, will be illustrated using the following PN situation:

PN: $A = 2$ vs. N + S: $A \neq 2$ (s known, $\alpha = .01$)

Data Sets: (See Tables 1 and 2)

1. $V_1, \ldots, V_{50}$ i.i.d. Pa(1;1)
2. $W_1, \ldots, W_{50}$ i.i.d. Pa(2;1)
3. $X_1, \ldots, X_{50}$ i.i.d. Pa(3;1)
4. $Y_1, \ldots, Y_{50}$ i.i.d. Pa(2;2)
5. $Z_1, \ldots, Z_{50}$ i.i.d. Pa(1;5)

<u>The Decision Rules are</u>:

1. Decide N + S iff $V(1) < 2$ or $V(1) > 2.193$. Since $V(1) = 1.010$ one decides N + S.
2. Decide N + S iff $W(1) < 2$ or $W(1) > 2.193$. Since $W(1) = 2.033$ one decides PN.

3. Decide $N+S$ iff $X(1) < 2$ or $X(1) > 2.193$. Since $X(1) = 3.001$ one decides $N+S$.
4. Decide $N+S$ iff $Y(1) < 2$ or $Y(1) > 2.094$. Since $Y(1) = 2.044$ one decides PN.
5. Decide $N+S$ iff $Z(1) < 2$ or $Z(1) > 2.037$. Since $Z(1) = 1.008$ one decides $N+S$.

<u>Example 6.2</u>: A unknown, $s = s_0$ known

<u>PN</u>: $L(A,s_0) \in \Omega(PRP)$ vs. <u>N + S</u>: $L(A,s_0) \notin \Omega(PRP)$

<u>Detector Statistics</u>: Four Kolmogorov-Smirnov statistics $(D_n^{(1)}, D_n^{(2)}, D_n^{(3)}, D_n^{(4)})$ and a Srinivasan-type statistic (D_n) are available.

<u>Decision Rules</u>: (1-4) Decide $N+S$ iff $D_n > d_{n-1,\alpha}$ where $d_{n-1,\alpha}$ is value from Kolmogorov-Smirnov table and $D_n = D_n^{(1)}, D_n^{(2)}, D_n^{(3)}, D_n^{(4)}$. (5) decide $N+S$ iff $\tilde{D}_n > (\frac{n-1}{n})d_{n-1,\alpha}$.

<u>Data Sets</u>: (see Tables 1 and 2)

1. $X_1, \ldots, X_{50}$ i.i.d. Pa(1;2)
2. $Y_1, \ldots, Y_{50}$ i.i.d. Pa(5;5)

	critical value (α=.01)	statistic value ($\underset{\sim}{X}$)	decide	statistic valus ($\underset{\sim}{Y}$)	decide
$D_{50}^{(1)}$	0.233	0.032	PN	0.109	PN
$D_{50}^{(2)}$	0.233	0.032	PN	0.109	PN
$D_{50}^{(3)}$	0.233	0.103	PN	0.064	PN
$D_{50}^{(4)}$	0.233	0.099	PN	0.055	PN
$\tilde{D}_{50}$	0.228	0.031	PN	0.107	PN

<u>Example 6.3</u>: $A = A_0$ known, s unknown, $L(A_0,s) \in \Omega(PRP)$

<u>PN</u>: $s = s_0$ vs. <u>N + S</u>: $s \neq s_0$

<u>Detector Statistic</u>: $T = 2s_0 \sum_{j=1}^{n} \ln \frac{X(j)}{A_0} \sim \chi^2_{2n}$

<u>Decision Rule</u>: Decide N + S iff

$$T > \chi^2_{2n,1-\alpha/2} \quad \text{or} \quad T < \chi^2_{2n,\alpha/2}$$

Test the following PN situation using generated Pareto data:

PN: $s = 2$ vs. N + S: $s \neq 2$ (A known, $\alpha = .01$)

<u>Data Sets</u>: (see Tables 1 and 2)

1. $X_1, \ldots, X_{50}$ i.i.d. Pa(1;2)
2. $Y_1, \ldots, Y_{50}$ i.i.d. Pa(5;5)

Decision Rule: Decide N + S iff $T > 140.2$ or $T < 82.4$

From data set 1: $T = 91.017$, so one decides PN.

From data set 2: $T = 48.737$, so one decides N + S.

Example 6.4: $A = A_0$ known, s unknown

PN: $L(A_0,s) \varepsilon \Omega(PRP)$ vs. N + S: $L(A_0,s) \notin \Omega(PRP)$

Detector Statistic: $D_n^* = \sup_{0<u<1} \left| \frac{1}{n-1} \sum_1^{n-1} \varepsilon\left(u - \frac{T_j}{T_n}\right) - u \right|$

where $T_r = \sum_{j=1}^{r} \ln\left(\frac{X_j}{A_0}\right)$

Decision Rule: Decide N + S iff $D_n^* > d_{n-1,\alpha}$ where $d_{n-1,\alpha}$ is appropriate value from Kolmogorov-Smirnov table.

Data Sets: (see Tables 1 and 2)

1. $X_1, \ldots, X_{50}$ i.i.d. Pa(1;2)
2. $Y_1, \ldots, Y_{50}$ i.i.d. Pa(5;5)

Decide N + S iff $D_{50}^* > .233$

Since $D_{50}^*(\underset{\sim}{X}) = 0.101$, one decides PN.

Since $D_{50}^*(\underset{\sim}{Y}) = 0.053$, one decides PN.

Example 6.5. A unknown, s unknown, $L(A,s) \in \Omega(PRP)$.

PN: $L(A,s) = L(A_0,s_0)$ vs. N + S: $L(A,s) \neq L(A_0,s_0)$

Detector Statistic $$T = -2 \sum_{j=1}^{n} \ln\left[\frac{A_0}{X(j)}\right]^{s_0} \sim X^2_{2n}$$

Decision Rule: Decide N + S iff

$$T > X^2_{2n,1-\alpha/2} \quad \text{or} \quad T < X^2_{2n,\alpha/2}$$

Test the following PN situation using generated Pareto data:

PN: $(A,s) = (1,2)$ vs. N + S: $(A,s) \neq (1,2)$

Data Set: (see Table 2)

1. $X_1, \ldots, X_{50}$ i.i.d. Pa(1;2)

Decide N + S iff $T > 140.2$ or $T < 82.4$

Since the calculated value of the test statistic is $T = 91.02$, one decides PN.

Example 6.6. A unknown s unknown, $L(A,s) \in \Omega(PRP)$

PN: $s \leq s_0$ vs. N + s: $s > s_0$

Detector Statistic: $$T = 2s_0\left[\sum_{j=1}^{n} \ln X_j - n \ln X(1)\right] \sim X^2_{2n-2,\alpha}$$

Decision Rule: Decide N + S iff $T < X^2_{2n-2,\alpha}$

Test the following PN situation using generated Pareto data:

$$\text{PN: } s \leq 3 \qquad \text{vs. } \quad N + S\text{: } s > 3 \qquad (\alpha = .05)$$

<u>Data Sets</u>: (see Tables 1 and 2)

1. $X_1, \ldots, X_{50}$ i.i.d. Pa(1;2)
2. $Y_1, \ldots, Y_{50}$ i.i.d. Pa(5;5)

The critical value is $X^2_{(98,.05)} = 76.5$

The calculated statistic values are

From data set 1: $T = 129.32$, so one decides PN.

From data set 2: $T = 72.64$, so one decides N + S.

<u>Example 6.7</u>. A unknown, s unknown $L(A,s) \in \Omega(PRP)$

$$\underline{\text{PN:}}\ A \leq A_0 \qquad \text{vs.} \quad \underline{N + S}\text{: } A > A_0$$

Detector Statistic: $$T = \frac{n(n-1)[\ln X(1) - \ln A_0]}{\sum_1^n \ln X_j - n \ln X(1)} \sim F_{(2,2n-2)}$$

<u>Decision Rule</u>: Decide N + S iff

i. $X(1) > A_0$ and

ii. $T > f_{(2,2n-2,1-\alpha)}$

Test the following PN situation using generated Pareto data:

$$\text{PN: } A \leq 3 \qquad \text{vs. } \quad N + S\text{: } A > 3$$

Data Sets: (see Tables I and 2)

1. $X_1, \ldots X_{50}$ i.i.d. Pa(1;2)
2. $Y_1, \ldots, Y_{50}$ i.i.d. Pa(5;5)

The critical value for T ($\alpha = .01$) is

$$f_{(2,98,.99)} = 4.87$$

Since $X(1) = 1.024$, one decides PN for data set 1.

Since $Y(1) = 5.006$, and $T_Y = 103.7$, one decides N + S for Data Set 2.

Example 6.8. A, s unknown

PN: $L(A,s) \varepsilon \Omega(PRP)$ vs. N + S: $L(A,s) \notin \Omega(PRP)$

Detector Statistics:

i. $$D_n^1 = \sup_z \left|\frac{1}{n-2} \sum_{j=1}^{n-2} \varepsilon\left(z - \frac{\mathcal{D}_j}{\mathcal{D}_{n-1}}\right) - z\right| \sim \text{K-S}(n-2)$$

ii. $$D_n^2 = \sup_z \left|\frac{1}{n-2} \sum_{j=1}^{n-2} \varepsilon\left(z - \frac{\mathcal{E}_j}{\mathcal{E}_{n-1}}\right) - z\right| \sim \text{K-S}(n - 2)$$

where
$$\mathcal{D}_r = \sum_{j=1}^{r-1} \ln\left[\frac{X(j+1)}{X(1)}\right] + (n - r)\ln\left[\frac{X(r+1)}{X(1)}\right] \qquad 1 \leq r \leq n - 1$$

and
$$\mathcal{E}_r = \sum_{j=1}^{r} (n + 1 - j) \ln\left[\frac{X(j + 1)}{X(j)}\right] \qquad 1 \leq r \leq n - 1$$

<u>Decision Rules</u>: Decide $N + S$ iff

i. $D_n^1 > d_{(n-2,\alpha)}$

ii. $D_n^2 > d_{(n-2,\alpha)}$

where $d(n - 2,\alpha)$ is appropriate value from Kolmogorov-Smirnov table.

<u>Data Sets</u>: (see Tables 1, 2 and 3)

1. $X_1, \ldots, X_{50}$ i.i.d. Pa(1;2)
2. $Y_1, \ldots, Y_{50}$ i.i.d. Pa(5;5)
3. $W_1, \ldots, W_{24}$ waiting times for "peaks" (i.e. R's) of heart-beat cycles.

The critical value for data sets 1 and 2 is:

$$d_{(48,.01)} = 0.22$$

The critical value for data set 3 is:

$$d_{(22,.01)} = 0.314$$

The calculated statistic values are:

$D_{50}^1(\underset{\sim}{X}) = 0.058$ Decide PN.

$D_{50}^2(\underset{\sim}{X}) = 0.058$ Decide PN.

$D_{50}^1(\underset{\sim}{Y}) = 0.042$ Decide PN.

$D_{50}^2(\underset{\sim}{Y}) = 0.043$ Decide PN.

$D_{24}^1(\underset{\sim}{W}) = 0.614$ Decide N+S.

$D_{24}^2(\underset{\sim}{W}) = 0.605$ Decide N+S.

Example 7.1. A_1, A_2 known; s_1, s_2 unknown

PN: $s_1 = s_2$ vs. N + S: $s_1 \neq s_2$

Detector Statistic: $$T = \frac{n \sum_{j=1}^{m} \ln(X_j/A_1)}{m \sum_{j=1}^{n} \ln(Y_j/A_2)} \sim F(2m,2n)$$

Decision Rule: Decide N + S iff

$$T > F_{(2m,2n,1-\alpha/2)} \quad \text{or} \quad T < F_{(2m,2n,\alpha/2)}$$

Data Sets: (See Tables 1 and 2)

1. $X_1, \ldots, X_{50}$ i.i.d. Pa(1;2)
2. $Y_1, \ldots, Y_{50}$ i.i.d. Pa(5;2)
3. $Z_1, \ldots, Z_{50}$ i.i.d. Pa(5;5)

The critical values are:

$$f_{(100,100,.995)} = 1.68$$

$$f_{(100,100,.005)} = 0.595$$

The calculated statistic values for each pair of data sets are given below:

From data sets 1 and 2: $T = 0.788$, so one decides PN.

From data sets 1 and 3: $T = 1.868$, so one decides N + S.

From data sets 2 and 3: $T = 0.422$, so one decides N + S.

Example 7.2. A_1, A_2 unknown; s_1, s_2 known

PN: $A_1 = A_2$ vs. N + S: $A_1 \neq A_2$

Detector Statistic: $T = \ln\left(\frac{Y(1)}{X(1)}\right) \sim F(x)$

where

$$F(x) = \begin{cases} 1 - \dfrac{ms_1}{(ms_1+ns_2)} e^{-ns_2 x} & x \geq 0 \\ \dfrac{ns_2}{(ms_1+ns_2)} e^{ms_1 x} & x < 0 \end{cases}$$

Decision Rule: Decide N + S iff

$$T > C_1 \quad \text{or} \quad T < C_2 \quad \text{where} \quad F(c_1) = 1 - \alpha/2$$

$$F(c_2) = \alpha/2$$

Data Sets: (see Tables 1 and 2)

1. $X_1, \ldots, X_{50}$ i.i.d. Pa(1;2)
2. $Y_1, \ldots, Y_{50}$ i.i.d. Pa(5;2)
3. $Z_1, \ldots, Z_{50}$ i.i.d. Pa(5;5)

The first two data sets constitute a special case $(s_1 = s_2 = s_0$ known) and will be treated in Example 7.2a. For the remaining two pairs of data sets, one has the following critical values: $(m = n = 50,$ $s_1 = 2,\ s_2 = 5,\ \alpha = .01)$

$$c_1 = 0.016 \quad \text{and} \quad c_2 = -0.050$$

The appropriate decision rule is:

Decide N + S iff $T > 0.016$ or $T < -0.050$.

The calculated statistic values are given below:

From data sets 1 and 3: $T = \ln(\frac{Z(1)}{X(1)}) = 1.587$, so one decides N + S.

From data sets 2 and 3: $T = \ln(\frac{Z(1)}{Y(1)}) = -0.008$, so one decides PN.

<u>Example 7.2a</u>. A_1, A_2 unknown; $s_1 = s_2 = s_0$ known

$$\underline{\text{PN}}: A_1 = A_2 \quad \text{vs.} \quad \underline{\text{N + S}}: A_1 \neq A_2$$

<u>Detector Statistic</u>: $T = \frac{Y(1)}{X(1)}$

<u>Decision Rule</u>: Decide N + S iff $T < b_1$ or $T > b_2$

where $b_1 = [\frac{N}{n}(\frac{\alpha}{2})]^{1/ms_0}$, $b_2 = [\frac{N}{m}(\frac{\alpha}{2})]^{-1/ns_0}$

<u>Data Sets</u>: (See Table 1)

1. $X_1, \ldots, X_{50}$ i.i.d. Pa(1;2)
2. $Y_1, \ldots, Y_{50}$ i.i.d. Pa(5;2)

The critical values, for $\alpha = .01$ are:

$$b_1 = 0.95 \qquad \text{and} \qquad b_2 = 1.05$$

The calculated statistic value is

$$T = \frac{Y(1)}{X(1)} = 4.929 \qquad \text{Decide N + S.}$$

Example 7.3. $A_1 = A_2 = A$ unknown, s_1, s_2 unknown

PN: $s_1 \le s_2$ vs. N + S: $s_1 > s_2$

Detector Statistic: $T = \frac{\eta_1}{\eta}$

where

$$\eta_1 = \sum_{j=1}^{m} \ln X_j - m \ln X(1)$$

and

$$\eta = \sum_{j=1}^{m} \ln X_j + \sum_{j=1}^{n} \ln Y_j - NW$$

where $W = \min\{\ln X(1), \ln Y(1)\}$

Decision Rule: Decide N + S iff $T < c$ where c is determined by

$$I_c(m - 1, n) = 1 - \alpha$$

Data Sets: (See Tables 1 and 2)

1. $X_1, \ldots, X_{50}$ i.i.d. Pa(5;2)
2. $Y_1, \ldots, Y_{50}$ i.i.d. Pa(5;5)

Computation of critical value:

$$I_c(49,50) = 0.99 \qquad (\text{take } \alpha = .01)$$

$c = 0.61$ (from tables of Incomplete Beta function)

The decision rule is:

Decide N + S iff $T < 0.61$

Since $T = \frac{\eta_1}{\eta} = \frac{28.396}{40.91} = .694$

one decides PN.

Example 7.4. A_1, A_2 unknown; $s_1 = s_2 = s$ unknown

PN: $A_1 \leq A_2$ vs. N + S: $A_1 > A_2$

Detector Statistic: $T = \ln X(1)$

Decision Rule: Decide N + S iff $T > c$ where $c = \frac{\eta c'}{n} + W$,

$$W = \min \{\ln X(1), \ln Y(1)\}$$

$$\eta = \sum_{j=1}^{m} \ln X_j + \sum_{j=1}^{n} \ln Y_j - NW$$

and c' is determined by

$$(\frac{n}{N})(1 - c')^{N-2} = \alpha$$

Data Sets: (See Tables 1 and 2)

1. $X_1, \ldots, X_{50}$ i.i.d. Pa(1;2)
2. $Y_1, \ldots, Y_{50}$ i.i.d. Pa(5;5)

Computation of critical value $(\alpha = .01)$:

$$(\frac{50}{100})(1 - c')^{98} = .01$$

$$c' = .04$$

since $\eta = 129.73$ and $W = \ln X(1) = .024$, one has

$$c = \frac{(129.73)(.04)}{50} + .024 = .128$$

since $T = .024$, one decides PN.

Example 7.5. A_1, A_2 unknown; s_1, s_2 unknown

PN: $s_1 = s_2$ vs. N + S: $s_1 \neq s_2$

Detector Statistic: $$T = \frac{(n-2)\eta_1}{(m-2)\eta_2} \sim F_{(2(m-1),\ 2(n-1))}$$

where $$\eta_1 = \sum_{j=1}^{m-1} \ln \frac{X(j+1)}{X(1)}$$

and
$$\eta_2 = \sum_{j=1}^{n-1} \ln \frac{Y(j+1)}{Y(1)}$$

Decision Rule: Decide N + S iff

$$T > f_{(2(m-1),2(n-1),\ (1-\alpha/2))}$$

or

$$T < f_{(2(m-1),2(n-1),\alpha/2)}$$

Data Sets: (see Tables 1 and 2)

1. $X_1, \ldots, X_{50}$ i.i.d. Pa(1;2)
2. $Y_1, \ldots, Y_{50}$ i.i.d. Pa(5;2)
3. $Z_1, \ldots, Z_{50}$ i.i.d. Pa(5;5)

Critical values:.

$$f_{(98,98,0.99)} = 1.69$$

$$f_{(98,98,0.01)} = 0.592$$

The decision rule is:

Decide $N + S$ iff $T > 1.69$ or $T < .592$

For data sets 1 and 2: $T = 0.760$, so one decides PN.

For data sets 1 and 3: $T = 1.779$, so one decides $N + S$.

For data sets 2 and 3: $T = 0.427$, so one decides $N + S$.

Example 7.6. No techniques for Case 7.6 have been developed by the authors.

Example 7.7. A_1, A_2 unknown; s_1, s_2 unknown

PN: $(A_1, s_1) = (A_2, s_2)$ vs. N + S: $(A_1, s_1) \neq (A_2, s_2)$

Detector Statistics

(i) $T_1 = \dfrac{(n - 2)\eta_1}{(m - 2)\eta_2} \sim F_{(2(m-1),2(n-1))}$

$(m = n)$ (ii) $T_2 = \dfrac{N(N-2)|\ln X(1) - \ln Y(1)|}{4(\eta_1 + \eta_2)} \sim F_{(2,2N-4)}$

Decision Rule: Decide $N + S$ iff

(i) $T_1 > f_{(2(m-1),2(n-1),1-\alpha/4)}$ or

$T_1 < f_{(2(m-1),2(n-1),\alpha/4)}$

(ii) If $f_{\alpha/4} \leq T_1 \leq f_{1-\alpha/4}$

decide $N + S$ iff

$$T_2 > f_{(2,2N-4,1-\alpha/2)}$$

<u>Data Sets</u>: (see Tables 1 and 2)

1. $X_1, \ldots, X_{50}$ i.i.d. Pa(1;2)
2. $Y_1, \ldots, Y_{50}$ i.i.d. Pa(5;2)
3. $Z_1, \ldots, Z_{50}$ i.i.d. Pa(5;5)

Critical values ($\alpha = .01$):

(i). $f_{(98,98,.9975)} = 1.77$

$f_{(98,98,.0025)} = .565$

(ii) $f_{(2,196,.995)} = 5.35$

Decide N + S iff

(i) $T_1 > 1.77$ or $T_1 < .565$

(ii) If $.565 \leq T_1 \leq 1.77$, decide N + S iff $T_2 > 5.35$

1. For $(\underset{\sim}{X}, \underset{\sim}{Y})$, $T_1 = 0.760$, $T_2 = 39.102$ so one decides N + S.
2. For $(\underset{\sim}{X}, \underset{\sim}{Z})$, $T_1 = 1.779$, $T_2 = 57.682$ so one decides N + S.
3. For $(\underset{\sim}{Y}, \underset{\sim}{Z})$, $T_1 = 0.427$, $T_2 = 0.251$ so one decides N + S.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 8-82	2. GOVT ACCESSION NO. AD-A120972	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) SIGNAL DETECTION FOR PARETO RENEWAL PROCESSES		5. TYPE OF REPORT & PERIOD COVERED TECHNICAL
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) C.B. BELL, R. AHMAD, C.J. PARK, R. LUI		8. CONTRACT OR GRANT NUMBER(s) N00014-80-C-0208
9. PERFORMING ORGANIZATION NAME AND ADDRESS SAN DIEGO STATE UNIVERSITY SAN DIEGO, CA 92182		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS OFFICE OF NAVAL RESEARCH ARLINGTON, VA 22217		12. REPORT DATE OCTOBER 82
		13. NUMBER OF PAGES 68
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Pareto distribution, renewal process, distribution-free; sufficient statistic; statistical noise; order statistics; Kolmogorov-Smirnov statistic; conditional distribution; unbiased, Lilliefor's statistic, Srinivasan statistic.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Employing minimal sufficient statistics maximal statistical noise, several Kolmogorov-type statistics; and conditional distributions, optimal detection procedures are constructed for various one-and-two-sample problems involving Pareto Renewal processes. Cases with and without nuisance parameters are treated. Optimal parametric and distribution-free procedures are developed.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)